U.S.S.N.: 10/521,300
Filed: 1/14/05
Case No.: 21126YP

BEST AVAILABLE COPY

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 July 2001 (05.07.2001)

PCT

(10) International Publication Number
WO 01/48209 A2

(51) International Patent Classification[7]: C12N 15/31, 15/11, 15/10, C07K 14/245

(21) International Application Number: PCT/US00/34419

(22) International Filing Date: 19 December 2000 (19.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/173,005 23 December 1999 (23.12.1999) US

(71) Applicant (*for all designated States except US*): **ELITRA PHARMACEUTICALS, INC.** [US/US]; Suite A, 3510 Dunhill Street, San Diego, CA 92121 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): **FORSYTH, R., Allyn** [US/US]; 1135 Beryl Street, San Diego, CA 92109 (US). **OHLSEN, Kari, L.** [US/US]; 3560 Vista De La Orilla, San Diego, CA 92117 (US). **ZYSKIND, Judith, W.** [US/US]; 8415 La Jolla Scenic Drive, La Jolla, CA 92037 (US).

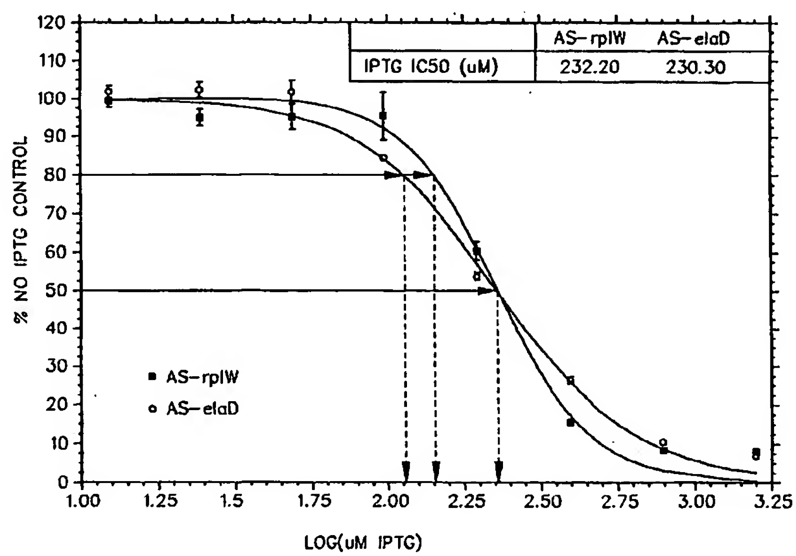
(74) Agent: **ISRAELSEN, Ned, A.**; 16th Floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION OF E. COLI



(57) **Abstract:** The sequences of nucleic acids encoding proteins required for E. coli proliferation are disclosed. The nucleic acids can also be used to screen for homologous genes that are required for proliferation in microorganisms other than E. coli. The nucleic acids can also be used to design expression vectors and secretion vectors. The nucleic acids can be used to express proteins or portions thereof, to obtain antibodies capable of specifically binding to the expressed proteins, and to use those expressed proteins as a screen to isolate candidate molecules for rational drug discovery programs. The nucleic acids of the present invention can also be used in various assay systems to screen for antimicrobial agents.

WO 01/48209 A2

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION OF E. COLI

BACKGROUND OF THE INVENTION

Since the discovery of penicillin, the use of antibiotics to treat the ravages of bacterial infections has saved millions of lives. With the advent of these "miracle drugs," for a time it was popularly believed that humanity might, once and for all, be saved from the scourge of bacterial infections. In fact, during the 1980s and early 1990s, many large pharmaceutical companies cut back or eliminated antibiotics research and development. They believed that infectious disease caused by bacteria finally had been conquered and that markets for new drugs were limited. Unfortunately, this belief was overly optimistic.

The tide is beginning to turn in favor of the bacteria as reports of drug resistant bacteria become more frequent. The United States Centers for Disease Control announced that one of the most powerful known antibiotics, vancomycin, was unable to treat an infection of the common *Staphylococcus aureus* (staph). This organism is commonly found in our environment and is responsible for many nosocomial infections. The import of this announcement becomes clear when one considers that vancomycin was used for years to treat infections caused by stubborn strains of bacteria, like staph. In short, the bacteria are becoming resistant to our most powerful antibiotics. If this trend continues, it is conceivable that we will return to a time when what are presently considered minor bacterial infections are fatal diseases.

There are a number of causes for the predicament in which practitioners of medical arts find themselves. Over-prescription and improper prescription habits by some physicians have caused an indiscriminate increase in the availability of antibiotics to the public. The patient is also partly responsible, for even in instances where an antibiotic is the appropriate treatment, patients will often improperly use the drug, the result being yet another population of bacteria that is resistant, in whole or in part, to traditional antibiotics.

The bacterial scourges that have haunted humanity remain, in spite of the development of modern scientific practices to deal with the diseases that they cause. Drug resistant bacteria are now advancing on the health of humanity. A new generation of antibiotics to once again deal with the pending health threat that bacteria present is required.

Discovery of New Antibiotics

As more and more bacterial strains become resistant to the panel of available antibiotics, new compounds are required. In the past, practitioners of pharmacology would have to rely upon traditional methods of drug discovery to generate novel, safe and efficacious compounds for the treatment of disease. Traditional drug discovery methods involve blindly testing potential drug candidate-molecules, often selected at random, in the hope that one might prove to be an effective treatment for some disease. The process is painstaking and laborious, with no guarantee of success. Today, the average cost to discover and develop a new drug is nearly US $500 million, and the

average time is 15 years from laboratory to patient. Improving this process, even incrementally, would represent a huge advance in the generation of novel antimicrobial agents.

Newly emerging practices in drug discovery utilize a number of biochemical techniques to provide for directed approaches to creating new drugs, rather than discovering them at random. For example, gene sequences and proteins encoded thereby that are required for the proliferation of an organism make for excellent targets since exposure of bacteria to compounds active against these targets would result in the inactivation of the organism. Once a target is identified, biochemical analysis of that target can be used to discover or to design molecules that interact with and alter the functions of the target. Using physical and computational techniques, to analyze structural and biochemical targets in order to derive compounds that interact with a target is called rational drug design and offers great future potential. Thus, emerging drug discovery practices use molecular modeling techniques, combinatorial chemistry approaches, and other means to produce and screen and/or design large numbers of candidate compounds.

Nevertheless, while this approach to drug discovery is clearly the way of the future, problems remain. For example, the initial step of identifying molecular targets for investigation can be an extremely time consuming task. It may also be difficult to design molecules that interact with the target by using computer modeling techniques. Furthermore, in cases where the function of the target is not known or is poorly understood, it may be difficult to design assays to detect molecules that interact with and alter the functions of the target. To improve the rate of novel drug discovery and development, methods of identifying important molecular targets in pathogenic microorganisms and methods for identifying molecules that interact with and alter the functions of such molecular targets are urgently required.

Escherichia coli represents an excellent model system to understand bacterial biochemistry and physiology. The estimated 4288 genes scattered along the 4.6×10^6 base pairs of the *Escherichia coli* (*E. coli*) chromosome offer tremendous promise for the understanding of bacterial biochemical processes. In turn, this knowledge will assist in the development of new tools for the diagnosis and treatment of bacteria-caused human disease. The entire *E. coli* genome has been sequenced, and this body of information holds a tremendous potential for application to the discovery and development of new antibiotic compounds. Yet, in spite of this accomplishment, the general functions or roles of many of these genes are still unknown. For example, the total number of proliferation-required genes contained within the *E. coli* genome is unknown, but has been variously estimated at around 200 to 700 (Armstrong, K.A. and Fan, D.P. Essential Genes in the *metB-malB* Region of *Escherichia coli* K12, 1975, J. Bacteriol. 126: 48-55).

Novel, safe and effective antimicrobial compounds are needed in view of the rapid rise of antibiotic resistant microorganisms. However, prior to this invention, the characterization of even a single bacterial gene was a painstaking process, requiring years of effort. Accordingly, there is an urgent need for more novel methods to identify and characterize bacterial genomic sequences that

encode gene products required for proliferation and for methods to identify molecules that interact with and alter the functions of such genes and gene products.

SUMMARY OF THE INVENTION

One embodiment of the present invention is a purified or isolated nucleic acid sequence consisting essentially of one of the sequence of nucleotides of SEQ ID NOs: 1-93, wherein expression of said nucleic acid in a microorganism is capable of inhibiting the proliferation of a microorganism. The nucleic acid sequence may have as sequence of nucleotides complementary to at least a portion of the nucleotide sequence of the coding strand of a gene whose expression is required for proliferation of a microorganism. The nucleic acid may have a nucleotide sequence complementary to at least a portion of the nucleotide sequence of an RNA required for proliferation of a microorganism. The nucleotide sequence of the RNA may encode more than one gene product.

Another embodiment of the present invention is a purified or isolated nucleic acid comprising a fragment of one of the nucleotide sequences of SEQ ID NOs.: 1-93, said fragment selected from the group consisting of fragments comprising at least 10, at least 20, at least 25, at least 30, at least 50 and more than 50 consecutive nucleotides of the nucleotide sequence of one of SEQ ID NOs: 1-93.

Another embodiment of the present invention is a vector comprising a promoter operably linked to the nucleic acid sequences of each of the preceding paragraphs. The promoter may be active in a microorganism selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling within the genera of any of the above species.

Another embodiment of the present invention is a host cell containing the vectors of the preceding paragraph.

Another embodiment of the present invention is a purified or isolated nucleic acid consisting essentially of the coding sequence of one of SEQ ID NOs: 106-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286.

Another embodiment of the present invention is a fragment of the nucleic acid of the preceding paragraph, said fragment comprising at least 10, at least 20, at least 25, at least 30, at least 50 or more than 50 consecutive nucleotides of one of SEQ ID NOs: 106-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286.

Another embodiment of the present invention is a vector comprising a promoter operably linked to the nucleic acid of the preceding two paragraphs.

Another embodiment of the present invention is a purified or isolated antisense nucleic acid comprising a nucleic acid sequence complementary to at least a portion of an intragenic sequence, intergenic sequence, sequences spanning at least a portion of two or more genes, 5' noncoding region, or 3' noncoding region within an operon comprising a proliferation-required gene whose activity or expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93.

Another embodiment of the present invention is a purified or isolated nucleic acid comprising a nucleic acid having at least 70% identity to a sequence selected from the group consisting of SEQ ID NOs.: 1-93, fragments comprising at least 25 consecutive nucleotides of SEQ ID NOs.: 1-93, the sequences complementary to SEQ ID NOs.: 1-93 and the sequences complementary to fragments comprising at least 25 consecutive nucleotides of SEQ ID NOs.: 1-93 as determined using BLASTN version 2.0 with the default parameters. The nucleic acid may be from an organism selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling within the genera of any of the above species.

Another embodiment of the present invention is a vector comprising a promoter operably linked to a nucleic acid encoding a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93. The polypeptide may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a host cell containing the vector of the preceding paragraph.

Another embodiment of the present invention is a purified or isolated polypeptide comprising a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93, or a fragment selected from the group consisting of fragments comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of one of the said polypeptides. The polypeptide may comprise a polypeptide comprising one of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 or a fragment comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a purified or isolated polypeptide comprising a polypeptide having at least 25% identity to a polypeptide whose expression is inhibited by a sequence selected from the group consisting of SEQ ID NOs.: 1-93, or at least 25% identity to a fragment comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide whose expression is inhibited by a nucleic acid selected from the group consisting of SEQ ID NOs.: 1-93 as determined using FASTA version 3.0t78 with the default parameters. The polypeptide may have at least 25% identity to a polypeptide comprising one of SEQ ID NOs: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 or at least 25% identity to a fragment comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide comprising one of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 as determined using FASTA version 3.0t78 with the default parameters.

Another embodiment of the present invention is an antibody capable of specifically binding one of the polypeptides of the preceding paragraph.

Another embodiment of the present invention is a method of producing a polypeptide, comprising introducing a vector comprising a promoter operably linked to a nucleic acid encoding a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93 into a cell and expressing said polypeptide. The method may further comprise the step of isolating said polypeptide. The polypeptide may comprise a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a method of inhibiting proliferation of a microorganism comprising inhibiting the activity or reducing the amount of a gene product whose expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group

consisting of SEQ ID NOs.: 1-93 or inhibiting the activity or reducing the amount of a nucleic acid encoding said gene product. The gene product may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a method for identifying a compound which influences the activity of a gene product required for proliferation, said gene product comprising a gene product whose expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising contacting said gene product with a candidate compound and determining whether said compound influences the activity of said gene product. The gene product may be a polypeptide and said activity may be an enzymatic activity. The gene product may be a polypeptide and said activity may be a carbon compound catabolism activity. The gene product may be a polypeptide and said activity may be a biosynthetic activity. The gene product may be a polypeptide and said activity may be a transporter activity. The gene product may be a polypeptide and said activity may be a transcriptional activity. The gene product may be a polypeptide and said activity may be a DNA replication activity. The gene product may be a polypeptide and said activity my be a cell division activity. The gene product may be a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a compound identified using the methods of the preceding paragraph.

Another embodiment of the present invention is a method for identifying a compound or nucleic acid having the ability to reduce the activity or level of a gene product required for proliferation, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising:

(a) providing a target that is a gene or RNA, wherein said target comprises a nucleic acid encoding said gene product;

(b) contacting said target with a candidate compound or nucleic acid; and

(c) measuring an activity of said target.

The target may be a messenger RNA molecule and said activity may be translation of said messenger RNA. The target may be a messenger RNA molecule and said activity may be transcription of a gene encoding said messenger RNA. The target may be a gene and said activity may be transcription of said gene. The target may be a nontranslated RNA and said activity may be processing or folding of said nontranslated RNA or assembly of said nontranslated RNA into a protein/RNA complex. The target gene or RNA may encode a polypeptide comprising a sequence

selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a compound or nucleic acid identified using the methods of the preceding paragraph.

Another embodiment of the present invention is a method for identifying a compound which reduces the activity or level of a gene product required for proliferation of a microorganism, wherein the activity or expression of said gene product is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising the steps of:

(a) expressing a sub-lethal level of an antisense nucleic acid complementary to a nucleic acid encoding said gene product in a cell to reduce the activity or amount of said gene product in said cell, thereby producing a sensitized cell;

(b) contacting said sensitized cell with a compound; and

(c) determining whether said compound inhibits the growth of said sensitized cell.

The determining step may comprise determining whether said compound inhibits the growth of said sensitized cell to a greater extent than said compound inhibits the growth of a nonsensitized cell. The cell may be selected from the group consisting of bacterial cells, fungal cells, plant cells, and animal cells. The cell may be a Gram negative bacterium. The cell may be an *E. coli* cell. The cell may be from an organism selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling within the genera of any of the above species. The antisense nucleic acid may be transcribed from an inducible promoter. The method may further comprise the step of contacting said cell with a concentration of inducer which induces said antisense nucleic acid to a sub-lethal level. Growth inhibition may be measured by monitoring optical density of a culture growth solution. The gene product may be a polypeptide. The polypeptide may comprise a sequence selected from the group consisting of SEQ ID NOs.: 299-

305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479. The gene product may be an RNA.

Another embodiment of the present invention is a compound identified using the methods of the preceding paragraph.

Another embodiment of the present invention is a method for inhibiting cellular proliferation comprising introducing a compound with activity against a gene whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 or a compound with activity against the product of said gene into a population of cells expressing said gene. The compound may be an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, or a proliferation-inhibiting portion thereof. The proliferation inhibiting portion of one of SEQ ID NOs.: 1-93 may be a fragment comprising at least 10, at least 20, at least 25, at least 30, at least 50 or more than 51 consecutive nucleotides of one of SEQ ID NOs.: 1-93. The population may be a population selected from the group consisting of bacterial cells, fungal cells, plant cells, and animal cells. The population may be a population of Gram negative bacteria. The population may be a population of *E. coli* cells. The population may be a population selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling within the genera of any of the above species. The gene may encode a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a preparation comprising an effective concentration of an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, or a proliferation-inhibiting portion thereof in a pharmaceutically acceptable carrier. The proliferation-inhibiting portion of one of SEQ ID NOs.: 1-93 may comprise at least 10, at least 20, at least 25, at least 30, at least 50 or more than 50 consecutive nucleotides of one of SEQ ID NOs.: 1-93.

Another embodiment of the present invention is a method for inhibiting the activity or expression of a gene in an operon required for proliferation wherein the activity or expression of at least one gene in said operon is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising contacting a cell in a cell population with an antisense nucleic acid comprising at least a proliferation-inhibiting portion of said operon. The antisense nucleic acid comprises a sequence selected from the group consisting of SEQ ID NOs.: 1-93 or a proliferation inhibiting portion thereof.

The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a plasmid which expresses said antisense nucleic acid into said cell population. The cell may be contacted with said antisense nucleic acid by introducing a phage which expresses said antisense nucleic acid into said cell population. The cell may be contacted with said antisense nucleic acid by expressing said antisense nucleic acid from the chromosome of cells in said cell population. The cell may be contacted with said antisense nucleic acid by introducing a promoter adjacent to a chromosomal copy of said antisense nucleic acid such that said promoter directs the synthesis of said antisense nucleic acid. The cell may be contacted with said antisense nucleic acid by introducing a retron which expresses said antisense nucleic acid into said cell population. The cell may be contacted with said antisense nucleic acid by introducing a ribozyme into said cell-population, wherein a binding portion of said ribozyme is complementary to said antisense oligonucleotide. The cell may be contacted with said antisense nucleic acid by introducing a liposome comprising said antisense oligonucleotide into said cell. The cell may be contacted with said antisense nucleic acid by electroporation of said antisense nucleic acid. The antisense nucleic acid may be a fragment comprising at least 10, at least 20, at least 25, at least 30, at least 50 or more than 50 consecutive nucleotides of one of SEQ ID NOs.: 1-93. The antisense nucleic acid may be an oligonucleotide.

Another embodiment of the present invention is a method for identifying a gene which is required for proliferation of a microorganism comprising:

(a) contacting a microorganism other than *E. coli* with a nucleic acid selected from the group consisting of SEQ ID NOs.: 1-93;

(b) determining whether said nucleic acid inhibits proliferation of said microorganism; and

(c) identifying the gene in said microorganism which is inhibited by said nucleic acid.

The microorganism may be a Gram negative bacterium. The microorganism may be selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans,*

Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis and any species falling within the genera of any of the above species. The method may further comprise introducing said nucleic acid into a vector functional in said microorganism prior to introducing said inhibitory nucleic acid into said microorganism.

Another embodiment of the present invention is a method for identifying a compound having the ability to inhibit proliferation of a microorganism comprising:

(a) identifying in a first microorganism a homolog of a gene or gene product present in a second microorganism which is different than said first microorganism, wherein the activity or level of said gene or gene product is inhibited by a nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs. 1-93 ;

(b) identifying an inhibitory nucleic acid sequence which inhibits the activity of said homolog in said first microorganism;

(c) contacting said first microorganism with a sub-lethal level of said inhibitory nucleic acid, thus sensitizing said first microorganism;

(d) contacting the sensitized microorganism of step (c) with a compound; and

(e) determining whether said compound inhibits proliferation of said sensitized microorganism.

The determining step may comprise determining whether said compound inhibits proliferation of said sensitized microorganism to a greater extent than said compound inhibits proliferation of a nonsensitized microorganism. Step (a) may comprise identifying a homologous nucleic acid to a gene or gene product whose activity or level is inhibited by a nucleic acid selected from the group consisting of SEQ ID NOs. 1-93 or a nucleic acid encoding a homologous polypeptide to a polypeptide whose activity or level is inhibited by a nucleic acid selected from the group consisting of SEQ ID NOs. 1-93 by using an algorithm selected from the group consisting of BLASTN version 2.0 with the default parameters and FASTA version 3.0t78 algorithm with the default parameters to identify said homologous nucleic acid or said nucleic acid encoding a homologous polypeptide in a database. Step (a) may comprise identifying a homologous nucleic acid or a nucleic acid encoding a homologous polypeptide by identifying nucleic acids which hybridize to said first gene. Step (a) may comprise expressing a nucleic acid selected from the group consisting of SEQ ID NOs. 1-93 in said microorganism. The inhibitory nucleic acid may be an antisense nucleic acid. The inhibitory nucleic acid may comprise an antisense nucleic acid to a portion of

said homolog. The inhibitory nucleic acid may comprise an antisense nucleic acid to a portion of the operon encoding said homolog. The step of contacting the first microorganism with a sub-lethal level of said inhibitory nucleic acid may comprise directly contacting said microorganism with said inhibitory nucleic acid. The step of contacting the first microorganism with a sub-lethal level of said inhibitory nucleic acid may comprise expressing an antisense nucleic acid to said homolog in said microorganism. The gene product may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a compound identified using the method of the preceding paragraph.

Another embodiment of the present invention is a method of identifying a compound having the ability to inhibit proliferation comprising:

(a) contacting a microorganism other than *E. coli* with a sub-lethal level of a nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs. 1-93 or a portion thereof which inhibits the proliferation of *E. coli*, thus sensitizing said microorganism;

(b) contacting the sensitized microorganism of step (a) with a compound; and

(c) determining whether said compound inhibits proliferation of said sensitized microorganism.

The determining step may comprise determining whether said compound inhibits proliferation of said sensitized microorganism to a greater extent than said compound inhibits proliferation of a nonsensitized microorganism.

Another embodiment of the present invention is a compound identified using the methods of the preceding paragraph.

Another embodiment of the present invention is a method for identifying a compound having activity against a biological pathway required for proliferation comprising:

(a) sensitizing a cell by expressing a sub-lethal level of an antisense nucleic acid complementary to a nucleic acid encoding a gene product required for proliferation, wherein the activity or expression of said gene product is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, in said cell to reduce the activity or amount of said gene product;

(b) contacting the sensitized cell with a compound; and

(c) determining whether said compound inhibits the growth of said sensitized cell.

The determining step may comprise determining whether said compound inhibits the growth of said sensitized cell to a greater extent than said compound inhibits the growth of a nonsensitized cell. The cell may be selected from the group consisting of bacterial cells, fungal cells, plant cells, and animal cells. The cell may be a Gram negative bacterium. The Gram negative bacterium may be *E. coli*. The cell may be selected from the group consisting of *Aspergillus fumigatus*, *Bacillus*

anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis and any species falling within the genera of any of the above species. The antisense nucleic acid may be transcribed from an inducible promoter. The method may further comprise contacting the cell with an agent which induces expression of said antisense nucleic acid from said inducible promoter, wherein said antisense nucleic acid is expressed at a sub-lethal level. The inhibition of proliferation may be measured by monitoring the optical density of a liquid culture. The gene product may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a compound identified using the methods of the preceding paragraph.

Another embodiment of the present invention is a method for identifying a compound having the ability to inhibit cellular proliferation comprising:

(a) contacting a cell with an agent which reduces the activity or level of a gene product required for proliferation of said cell, wherein said gene product is a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93;

(b) contacting said cell with a compound; and

(c) determining whether said compound reduces proliferation of said contacted cell.

The determining step may comprise determining whether said compound reduces proliferation of said contacted cell to a greater extent than said compound reduces proliferation of cells which have not been contacted with said agent. The agent which reduces the activity or level of a gene product required for proliferation of said cell may comprise an antisense nucleic acid to a gene or operon required for proliferation. The agent which reduces the activity or level of a gene product required for proliferation of said cell may comprise a compound known to inhibit growth or proliferation of a microorganism. The cell may contain a mutation which reduces the activity or level of said gene product required for proliferation of said cell. The mutation may be a temperature sensitive

mutation. The gene product may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a compound identified using the method of the preceding paragraph.

Another embodiment of the present invention is a method for identifying the biological pathway in which a proliferation-required gene or its gene product lies, wherein said gene or gene product comprises a gene or gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising:

(a) expressing a sub-lethal level of an antisense nucleic acid which inhibits the activity of said proliferation-required gene or gene product in a cell;

(b) contacting said cell with a compound known to inhibit growth or proliferation of a microorganism, wherein the biological pathway on which said compound acts is known; and

(c) determining whether said cell is sensitive to said compound.

The determining step may comprise determining whether said cell has a substantially greater sensitivity to said compound than a cell which does not express said sub-lethal level of said antisense nucleic acid and wherein said gene or gene product lies in the same pathway on which said compound acts if said cell expressing said sub-lethal level of said antisense nucleic acid has a substantially greater sensitivity to said compound than said cell which does not express said sub-lethal level of said antisense nucleic acid.

The gene product may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

Another embodiment of the present invention is a method for determining the biological pathway on which a test compound acts comprising:

(a) expressing a sub-lethal level of an antisense nucleic acid complementary to a proliferation-required nucleic acid in a cell, wherein the activity or expression of said proliferation-required nucleic acid is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 and wherein the biological pathway in which said proliferation-required nucleic acid or a protein encoded by said proliferation-required polypeptide lies is known,

(b) contacting said cell with said test compound; and

(c) determining whether said cell is sensitive to said test compound.

The determining step may comprise determining whether said cell has a substantially greater sensitivity to said test compound than a cell which does not express said sub-lethal level of said antisense nucleic acid. The method may further comprise:

(d) expressing a sub-lethal level of a second antisense nucleic acid complementary to a second proliferation-required nucleic acid in a second cell, wherein said second proliferation-required nucleic acid is in a different biological pathway than said proliferation-required nucleic acid in step (a); and

(e) determining whether said second cell does not have a substantially greater sensitivity to said test compound than a cell which does not express said sub-lethal level of said second antisense nucleic acid, wherein said test compound is specific for the biological pathway against which the antisense nucleic acid of step (a) acts if said second cell does not have substantially greater sensitivity to said test compound.

Another embodiment of the present invention is a purified or isolated nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93.

Another embodiment of the present invention is a compound which interacts with a gene or gene product whose activity or expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93 to inhibit proliferation.

Another embodiment of the present invention is a compound which interacts with a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93 to inhibit proliferation.

Another embodiment of the present invention is a method for manufacturing an antibiotic comprising the steps of screening one or more candidate compounds to identify a compound that reduces the activity or level of a gene product required for proliferation, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 and manufacturing the compound so identified.

The screening step may comprise performing any one of the methods of identifying a compound described above.

Another embodiment of the present invention is a method for inhibiting proliferation of a microorganism in a subject comprising administering a compound that reduces the activity or level of a gene product required for proliferation of said microorganism, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 to said subject. The method of subject may be selected from the group consisting of vertebrates, mammals, avians, and human beings. The gene product may comprise a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is an IPTG dose response curve in *E. coli* transformed with an IPTG-inducible plasmid containing either an antisense clone to the *E. coli rplW* gene (AS-*rplW*) which encodes a ribosomal protein required for protein synthesis and essential for cell proliferation, or an antisense clone to the *elaD* gene (AS-*elaD*) which is not known to be involved in protein synthesis and which is also essential for proliferation.

Figure 2A is a tetracycline dose response curve in *E. coli* transformed with an IPTG-inducible plasmid containing antisense to the *rplW* gene (AS-*rplW*) which was carried out in the presence of 0, 20 or 50 μM IPTG.

Figure 2B is a tetracycline dose response curve in *E. coli* transformed with an IPTG-inducible plasmid containing antisense to the *elaD* gene (AS-*elaD*) which was carried out in the presence of 0, 20 or 50 μM IPTG.

Figure 3 is a graph showing the fold increase in tetracycline sensitivity of *E. coli* transfected with antisense clones to essential ribosomal protein genes *L23* (AS-*rplW*) and *L7/L12* and *L10* (AS-*rplLrplJ*). Antisense clones to genes known not to be involved in protein synthesis (*atpB/E*(AS-*atpB/E*), *visC* (AS-*visC*), *elaD* (AS-*elaD*), *yohH* (AS-*yohH*) are much less sensitive to tetracycline.

Definitions

By "biological pathway" is meant any discrete cell function or process that is carried out by a gene product or a subset of gene products. Biological pathways include enzymatic, biochemical and metabolic pathways as well as pathways involved in the production of cellular structures such as cell walls. Biological pathways that are usually required for proliferation of microorganisms include, but are not limited to, cell division, DNA synthesis and replication, RNA synthesis (transcription), protein synthesis (translation), protein processing, protein transport, fatty acid biosynthesis, cell wall synthesis, cell membrane production, synthesis and maintenance, and the like.

By "inhibit activity of a gene or gene product" is meant having the ability to interfere with the function of a gene or gene product in such a way as to decrease expression of the gene or to reduce the level or activity of a product of the gene. Agents which inhibit the activity of a gene include agents that inhibit transcription of the gene, agents that inhibit processing of the transcript of the gene, agents that reduce the stability of the transcript of the gene, and agents that inhibit translation of the mRNA transcribed from the gene. In microorganisms, agents which inhibit the activity of a gene can act to decrease expression of the operon in which the gene resides or alter the folding or processing of operon RNA so as to reduce the level or activity of the gene product. The gene product can be a non-translated RNA such as ribosomal RNA, a translated RNA (mRNA) or

the protein product resulting from translation of the gene mRNA. Of particular utility to the present invention are antisense RNAs that have activities against the operons or genes to which they specifically hybridze.

By "activity against a gene product" is meant having the ability to inhibit the function or to reduce the level or activity of the gene product in a cell.

By "activity against a protein" is meant having the ability to inhibit the function or to reduce the level or activity of the protein in a cell.

By "activity against a nucleic acid" is meant having the ability to inhibit the function or to reduce the level or activity of the nucleic acid in a cell.

By "activity against a gene" is meant having the ability to inhibit the function or expression of the gene in a cell.

By "activity against an operon" is meant having the ability to inhibit the function or reduce the level of one or more products of the operon in a cell.

By "antibiotic" is meant an agent which inhibits the proliferation of a microorganism.

By "*E. coli or Escherichia coli*" is meant *Escherichia coli* or any organism previously categorized as a species of *Shigella* including *Shigella boydii, Shigella flexneri, Shigella dysenteriae, Shigella sonnei, Shigella 2A.*

By "identifying a compound" is meant to screen one or more compounds in a collection of compounds such as a combinatorial chemical library or other library of chemical compounds or to characterize a single compound by testing the compound in a given assay and determining whether it exhibits the desired activity.

By "inducer" is meant an agent or solution which, when placed in contact with a microorganism, increases transcription from a desired promoter.

As used herein, "nucleic acid" means DNA, RNA, or modified nucleic acids. Thus, the terminology "the nucleic acid of SEQ ID NO: X" includes both the DNA sequence of SEQ ID NO: X and an RNA sequence in which the thymidines in the DNA sequence have been substituted with uridines in the RNA sequence and in which the deoxyribose backbone of the DNA sequence has been substituted with a ribose backbone in the RNA sequence. Modified nucleic acids are nucleic acids having nucleotides or structures which do not occur in nature, such as nucleic acids in which the internucleotide phosphate residues with methylphosphonates, phosphorothioates, phosphoramidates, and phosphate esters. Nonphosphate internucleotide analogs such as siloxane bridges, carbonate brides, thioester bridges, as well as many others known in the art may also be used in modified nucleic acids. Modified nucleic acids may also comprise, α-anomeric nucleotide units and modified nucleotides such as 1,2-dideoxy-d-ribofuranose, 1,2-dideoxy-1-phenylribofuranose, and N^4, N^4-ethano-5-methyl-cytosine are contemplated for use in the present invention. Modified nucleic acids may also be peptide nucleic acids in which the entire

deoxyribose-phosphate backbone has been exchanged with a chemically completely different, but structurally homologous, polyamide (peptide) backbone containing 2-aminoethyl glycine units.

As used herein, "sub-lethal" means a concentration of an agent below the concentration required to inhibit all cell growth.

DETAILED DESCRIPTION OF THE INVENTION

The present invention describes a group of *E. coli* genes and gene families required for growth and/or proliferation. A proliferation-required gene or gene family is one where, in the absence of a gene transcript and/or gene product, growth or viability of the microorganism is reduced or eliminated. Thus, as used herein the terminology "proliferation-required" or "required for proliferation" encompasses instances where the absence of a gene transcript and/or gene product completely eliminates cell growth as well as instances where the absence of a gene transcript and/or gene product merely reduces cell growth. These proliferation-required genes can be used as potential targets for the generation of new antimicrobial agents. To achieve that goal, the present invention also encompasses novel assays for analyzing proliferation-required genes and for identifying compounds which interact with the gene products of the proliferation-required genes. In addition, the present invention contemplates the expression of genes and the purification of the proteins encoded by the nucleic acid sequences identified as required proliferation genes and reported herein. The purified proteins can be used to generate reagents and screen small molecule libraries or other candidate compound libraries for compounds that can be further developed to yield novel antimicrobial compounds. The present invention also describes methods for identification of homologous genes or polypeptides in organisms other than *E. coli*.

The present invention utilizes a novel method to identify proliferation-required *E. coli* sequences. Generally, a library of nucleic acid sequences from a given source are subcloned or otherwise inserted into an inducible expression vector, thus forming an expression library. Although the insert nucleic acids may be derived from the chromosome of the organism into which the expression vector is to be introduced, because the insert is not in its natural chromosomal location, the insert nucleic acid is an exogenous nucleic acid for the purposes of the discussion herein. The term expression is defined as the production of an RNA molecule from a gene, gene fragment, genomic fragment, or operon. Expression can also be used to refer to the process of peptide or polypeptide synthesis. An expression vector is defined as a vehicle by which a ribonucleic acid (RNA) sequence is transcribed from a nucleic acid sequence carried within the expression vehicle. The expression vector can also contain features that permit translation of a protein product from the transcribed RNA message expressed from the exogenous nucleic acid sequence carried by the expression vector. Accordingly, an expression vector can produce an RNA molecule as its sole product or the expression vector can produce a RNA molecule that is ultimately translated into a protein product.

Once generated, the expression library containing the exogenous nucleic acid sequences is introduced into an *E. coli* population to search for genes that are required for bacterial proliferation.

Because the library molecules are foreign to the population of *E. coli*, the expression vectors and the nucleic acid segments contained therein are considered exogenous nucleic acid.

Expression of the exogenous nucleic acid fragments in the test population of *E. coli* containing the expression vector library is then activated. Activation of the expression vectors consists of subjecting the cells containing the vectors to conditions that result in the expression of the exogenous nucleic acid sequences carried by the expression vector library. The test population of *E. coli* cells is then assayed to determine the effect of expressing the exogenous nucleic acid fragments on the test population of cells. Those expression vectors that, upon activation and expression, negatively impact the growth of the *E. coli* screen population are identified, isolated, and purified for further study.

A variety of assays are contemplated to identify nucleic acid sequences that negatively impact growth upon expression. In one embodiment, growth in *E. coli* cultures expressing exogenous nucleic acid sequences is compared to growth in cultures not expressing these sequences. Optical density is used to monitor the extent of growth. Alternatively, enzymatic assays can be used to determine bacterial growth rates to identify exogenous nucleic acid sequences of interest. Colony size, colony morphology, and cell morphology are additional factors used to evaluate growth of the host cells. Those cultures that fail to grow or grow at a reduced rate under expression conditions are identified as containing an expression vector encoding a nucleic acid fragment that negatively affects a proliferation-required gene.

Once exogenous nucleic acid sequences of interest are identified, they are analyzed. The first step of the analysis is to acquire the nucleic acid sequence of the nucleic acid fragment of interest. To achieve this end, the insert in those expression vectors identified as containing a sequence of interest is sequenced, using standard techniques well known in the art. The next step of the process is to determine the source of the nucleic acid sequence.

Determination of sequence source is achieved by comparing the obtained sequence data with known sequences in various genetic databases. The sequences identified are used to probe these gene databases. The result of this procedure is a list of exogenous nucleic acid sequences corresponding to a list that includes novel bacterial genes required for proliferation as well as genes previously identified as required for proliferation.

The number of DNA and protein sequences available in database systems has been growing exponentially for years. For example, at the end of 1998, the complete sequences of *Caenorhabditis elegans*, *Saccharomyces cerevisiae* and nineteen bacterial genomes, including *E. coli* were available. This sequence information is stored in a number of databanks, such as GenBank (the National Center for Biotechnology Information (NCBI), and is publicly available for searching.

A variety of computer programs are available to assist in the analysis of the sequences stored within these databases. FASTA, (W. R. Pearson (1990) "Rapid and Sensitive Sequence Comparison with FASTP and FASTA" Methods in Enzymology 183:63- 98), Sequence Retrieval System (SRS), (Etzold & Argos, SRS an indexing and retrieval tool for flat file data libraries. Comput. Appl.

Biosci. 9:49-57, 1993) are two examples of computer programs that can be used to analyze sequences of interest. In one embodiment of the present invention, the BLAST family of computer programs, which includes BLASTN version 2.0 with the default parameters, or BLASTX version 2.0 with the default parameters, is used to analyze nucleic acid sequences. BLAST, an acronym for "Basic Local Alignment Search Tool," is a family of programs for database similarity searching. The BLAST family of programs includes: BLASTN, a nucleotide sequence database searching program, BLASTX, a protein database searching program where the input is a nucleic acid sequence; and BLASTP, a protein database searching program where the input is an amino acid sequence. BLAST programs embody a fast algorithm for sequence matching, rigorous statistical methods for judging the significance of matches, and various options for tailoring the program for special situations. Assistance in using the program can be obtained by e-mail at blast@ncbi.nlm.nih.gov.

Bacterial genes are often transcribed in polycistronic groups. These groups comprise operons, which are a collection of genes and intergenic sequences. The genes of an operon are co-transcribed and often have related functions. Given the nature of the screening protocol, it is possible that the identified exogenous nucleic acid sequence corresponds to a gene or portion thereof with or without adjacent noncoding sequences, an intragenic sequence (i.e. a sequence within a gene), an intergenic sequence (i.e. a sequence between genes), a sequence spanning at least a portion of two or more genes, a 5' noncoding region or a 3' noncoding region located upstream or downstream from the actual sequence that is required for bacterial proliferation. Accordingly, determining which gene(s) that is encoded within the operons is individually required for proliferation is often desirable.

In one embodiment of the present invention, an operon is dissected to determine which gene or genes are required for proliferation. For example, the RegulonDB DataBase described by Huerta et al. (*Nucl. Acids Res.* 26:55-59, 1998), which may also be found on the website http://www.cifn.unam.mx/Computational_Biology/regulondb/,may be used to identify the boundaries of operons encoded within microbial genomes. A number of techniques that are well known in the art can then be used to dissect the operon. In one aspect of this embodiment, gene disruption by homologous recombination is used to individually inactivate the genes of an operon that is thought to contain a gene required for proliferation.

Several gene disruption techniques have been described for the replacement of a functional gene with a mutated, non-functional (null) allele. These techniques generally involve the use of homologous recombination. The method described by Link et al. (J. Bacteriol 1997 179:6228) serves as an excellent example of these methods as applicable to disruption of genes in *E. coli*. This technique uses crossover PCR to create a null allele with an in-frame deletion of the coding region of a target gene. The null allele is constructed in such a way that sequences adjacent to the wild type gene (ca. 500 bp) are retained. These homologous sequences surrounding the deletion null

allele provide targets for homologous recombination so that the wild type gene on the *E. coli* chromosome can be replaced by the constructed null allele.

The crossover PCR amplification product is subcloned into the vector pKO3, the features of which include a chloramphenicol resistance gene, the counter-selectable marker *sacB*, and a temperature sensitive autonomous replication function. Following transformation of an *E. coli* cell population with such a vector, selection for cells that have undergone homologous recombination of the vector into the chromosome is achieved by growth on chloramphenicol at the non-permissive temperature of 43°C. Under these conditions, autonomous replication of the plasmid cannot occur and cells are resistant to chloramphinicol only if the chloramphenicol resistance gene has been integrated into the chromosome. Usually a single crossover event is responsible for this integration event such that the *E. coli* chromosome now contains a tandem duplication of the target gene consisting of one wild type allele and one deletion null allele separated by vector sequence.

This new *E. coli* strain containing the tandem duplication can be maintained at permissive temperatures in the presence of drug selection (chloramphenicol). Subsequently, cells of this new strain are cultured at the permissive temperature 30°C without drug selection. Under these conditions, the chromosome of some of the cells within the population will have undergone an internal homologous recombination event resulting in removal of the plasmid sequences. Subsequent culturing of the strain in growth medium lacking chloramphenicol but containing sucrose is used to select for such recombinative resolutions. In the presence of the counter-selectable marker *sacB*, sucrose is rendered into a toxic metabolite. Thus, cells that survive this counter-selection have lost both the plasmid sequences from the chromosome and the autonomously replicating plasmid that results as a byproduct of recombinative resolution.

There are two possible outcomes of the above recombinative resolution via homologous recombination. Either the wild type copy of the targeted gene is retained on the chromosome or the mutated null allele is retained on the chromosome. In the case of an essential gene, a single copy of the null allele would be lethal and such cells should not be obtained by the above procedure when applied to essential genes. In the case of a non-essential gene, roughly equal numbers of cells containing null alleles and cells containing wild type alleles should be obtained. Thus, the method serves as a test for essentiality of the targeted gene: when applied to essential genes, only cells with a wild type allele on the chromosome will be obtained.

Other techniques have also been described for the creation of disruption mutations in *E. coli*. For example, Link et al. also describe inserting an in-frame sequence tag concommitantly with an in-frame deletion in order to simplify analysis of recombinants obtained. Further, Link et al. describe disruption of genes with a drug resistance marker such as a kanamycin resistance gene. Arigoni et al., (Arigoni, F. et al. A Genome-based Approach for the Identification of Essential Bacterial Genes, Nature Biotechnology 16: 851-856) describe the use of gene disruption combined with engineering a second copy of a test gene such that the expression of the gene is regulated by

and inducible promoter such as the arabinose promoter to test the essentiality of the gene. Many of these techniques result in the insertion of large fragments of DNA into the gene of interest, such as a drug selection marker. An advantage of the technique described by Link et al. is that it does not rely on an insertion into the gene to cause a functional defect, but rather results in the precise removal of the coding region. This insures the lack of polar effects on the expression of genes downstream from the target gene.

Recombinant DNA techniques can be used to express the entire coding sequences of the gene identified as required for proliferation, or portions thereof. The over-expressed proteins can be used as reagents for further study. The identified exogenous sequences are isolated, purified, and cloned into a suitable expression vector using methods well known in the art. If desired, the nucleic acids can contain the sequences encoding a signal peptide to facilitate secretion of the expressed protein.

Expression of fragments of the bacterial genes identified as required for proliferation is also contemplated by the present invention. The fragments of the identified genes can encode a polypeptide comprising at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 55, at least 60, at least 65, at least 75, or more than 75 consecutive amino acids of a gene complementary to one of the identified sequences of the present invention. The nucleic acids inserted into the expression vectors can also contain sequences upstream and downstream of the coding sequence.

When expressing the coding sequence of an entire gene identified as required for bacterial proliferation or a fragment thereof, the nucleic acid sequence to be expressed is operably linked to a promoter in an expression vector using conventional cloning technology. The expression vector can be any of the bacterial, insect, yeast, or mammalian expression systems known in the art. Commercially available vectors and expression systems are available from a variety of suppliers including Genetics Institute (Cambridge, MA), Stratagene (La Jolla, California), Promega (Madison, Wisconsin), and Invitrogen (San Diego, California). If desired, to enhance expression and facilitate proper protein folding, the codon usage and codon bias of the sequence can be optimized for the particular expression organism in which the expression vector is introduced, as explained by Hatfield, et al., U.S. Patent No. 5,082,767. Fusion protein expression systems are also contemplated by the present invention.

Following expression of the protein encoded by the identified exogenous nucleic acid sequence, the protein is purified. Protein purification techniques are well known in the art. Proteins encoded and expressed from identified exogenous nucleic acid sequences can be partially purified using precipitation techniques, such as precipitation with polyethylene glycol. Alternatively, epitope tagging of the protein can be used to allow simple one step purification of the protein. Chromatographic methods usable with the present invention can include ion-exchange chromatography, gel filtration, use of hydroxyapaptite columns, immobilized reactive dyes, chromatofocusing, and use of high-performance liquid chromatography. Electrophoretic methods such one-dimensional gel electrophoresis, high-resolution two-dimensional polyacrylamide electrophoresis,

isoelectric focusing, and others are contemplated as purification methods. Also, affinity chromatographic methods, comprising antibody columns, ligand presenting columns and other affinity chromatographic matrices are contemplated as purification methods in the present invention.

The purified proteins produced from the gene coding sequences identified as required for proliferation can be used in a variety of protocols to generate useful antimicrobial reagents. In one embodiment of the present invention, antibodies are generated against the proteins expressed from the identified exogenous nucleic acid sequences. Both monoclonal and polyclonal antibodies can be generated against the expressed proteins. Methods for generating monoclonal and polyclonal antibodies are well known in the art. Also, antibody fragment preparations prepared from the produced antibodies discussed above are contemplated.

In addition, the purified protein, fragments therof, or derivatives thereof may be administered to an individual in a pharmaceutically acceptable carrier to induce an immune response against the protein. Preferably, the immune response is a protective immune response which protects the individual. Methods for determining appropriate dosages of the protein and pharmaceutically acceptable carriers are familiar to those skilled in the art.

Another application for the purified proteins of the present invention is to screen small molecule libraries for candidate compounds active against the various target proteins of the present invention. Advances in the field of combinatorial chemistry provide methods, well known in the art, to produce large numbers of candidate compounds that can have a binding, or otherwise inhibitory effect on a target protein. Accordingly, the screening of small molecule libraries for compounds with binding affinity or inhibitory activity for a target protein produced from an identified gene sequence is contemplated by the present invention.

The present invention further contemplates utility against a variety of other pathogenic organisms in addition to *E. coli*. For example, the invention has utility in identifying genes required for proliferation in prokaryotes and eukaryotes. For example, the invention has utility with protists, such as *Plasmodium* spp .and as *Entamoeba* spp.; plants; animals, such *and Contracaecum* spp; and fungi including *Candida* spp., ((e.g., *Candida albicansCandida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis,*)), *Saccharomyces cerevisiae, Cryptococcus neoformans*, and *Aspergillus fumigatus*. In one embodiment of the present invention, monera, specifically bacteria are probed in search of novel gene sequences required for proliferation. This embodiment is particularly important given the rise of drug resistant bacteria.

The numbers of bacterial species that are becoming resistant to existing antibiotics are growing. A partial list of these organisms includes: *Staphylococcus* spp., such as *S. aureus*; *Enterococcus* spp., such as *E. faecalis*; *Pseudomonas* spp., *such as P. aeruginosa, Clostridium* spp., such as *C. botulinum or C. difficile, Haemophilus* spp., such as *H. influenzae, Enterobacter* spp., such as *E. cloacae, Vibrio* spp., such as *V. cholera; Moraxala* spp., such as *M. catarrhalis*;

Streptococcus spp., such as *S. pneumoniae, Neisseria* spp., such as *N. gonorrhoeae; Mycoplasma* spp., such as *Mycoplasma pneumoniae; Salmonella typhimurium*; *Helicobacter pylori*; *Escherichia coli;* and *Mycobacterium tuberculosis*. The sequences identified as required for proliferation in the present invention can be used to probe these and other organisms to identify homologous required proliferation genes contained therein.

In one embodiment of the present invention, the nucleic acid sequences disclosed herein are used to screen genomic libraries generated from bacterial species of interest other than *E. coli*. For example, the genomic library may be from *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* or any species falling within the genera of any of the above species. Standard molecular biology techniques are used to generate genomic libraries from various microorganisms. In one aspect, the libraries are generated and bound to nitrocellulose paper. The identified exogenous nucleic acid sequences of the present invention can then be used as probes to screen the libraries for homologous sequences. The homologous sequences identified can then be used as targets for the identification of new, antimicrobial compounds with activity against more than one organism.

For example, the preceding methods may be used to isolate nucleic acids having a sequence with at least 97%, at least 95%, at least 90%, at least 85%, at least 80%, or at least 70% identity to a nucleic acid sequence selected from the group consisting of one of the sequences of SEQ ID NOS. 1-93, 106-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286, fragments comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides thereof, and the sequences complementary thereto. Identity may be measured using BLASTN version 2.0 with the default parameters. (Altschul, S.F. et al. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs, Nucleic Acid Res. 25: 3389-3402 (1997)). For example, the homologous polynucleotides may have a coding sequence which is a naturally occurring allelic variant of one of the coding sequences described herein. Such allelic variants may have a substitution, deletion or addition of one or more nucleotides when compared to

the nucleic acids of SEQ ID NOs: 1-93, 106-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286 or the sequences complementary thereto.

Additionally, the above procedures may be used to isolate nucleic acids which encode polypeptides having at least 99%, 95%, at least 90%, at least 85%, at least 80%, at least 70%, at least 60%, at least 50%, or at least 40% identity or similarity to a polypeptide having the sequence of one of SEQ ID NOs: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 or to a polypeptide whose expression is inhibited by a nucleic acid of one of SEQ ID NOs.: 1-93, or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids of the preceding polypeptides as determined using the FASTA version 3.0t78 algorithm with the default parameters. Alternatively, protein identity or similarity may be identified using BLASTP with the default parameters, BLASTX with the default parameters, or TBLASTN with the default parameters. (Alschul, S.F. et al. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs, Nucleic Acid Res. 25: 3389-3402 (1997)).

Alternatively, homologous nucleic acids or polypeptides may be identified by searching a database to identify sequences having a desired level of homology to a nucleic acid or a polypeptide involved in proliferation or an antisense nucleic acid to a nucleic acid involved in microbial proliferation. A variety of such databases are available to those skilled in the art, including GenBank and GenSeq. In some embodiments, the databases are screened to identify nucleic acids or polypeptides having at least 97%, at least 95%, at least 90%, at least 85%, at least 80%, at least 70%, at least 60%, or at least 50%, at least 40% identity or similarity to a nucleic acid or polypeptide involved in proliferation or an antisense nucleic acid involved in proliferation. For example, the database may be screened to identify nucleic acids homologous to one of SEQ ID Nos. 1-93, 106-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286, homologous to fragments comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides thereof, or homologous to the sequences complementary to any of the preceding nucleic acids. In other embodiments, the databases are screened to indetify polypeptides having at least 99%, 95%, 90%, 855, 80%, 70%, 60%, 50%, 40%, or at least 25% identity or similarity ot a polypeptide involved in proliferation or a portion thereof. For example, the database may be screened to identify polypeptides homologous to a polypeptide comprising one of SEQ ID NOs. 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479, a polypeptide whose expression is inhibited by a nucleic acid of one of SEQ ID NOs: 1-93, or homologous to fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids of any of the preceding polypeptides. In some embodiments, the database may be screened to identify homologous nucleic acids or polypeptides from organisms other than *E. coli*, including organisms such as *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr*

(also called Candida pseudotropicalis), Candida dubliniensis, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis,Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulimum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis or any species falling within the genera of any of the above species.

In another embodiment, gene expression arrays and microarrays can be employed. Gene expression arrays are high density arrays of DNA samples deposited at specific locations on a glass chip, nylon membrane, or the like. Such arrays can be used by researchers to quantify relative gene expression under different conditions. Gene expression arrays are used by researchers to help identify optimal drug targets, profile new compounds, and determine disease pathways. An example of this technology is found in U.S. Patent No. 5807522.

It is possible to study the expression of all genes in the genome of a particular microbial organism using a single array. For example, the arrays from Genosys consist of 12 x 24 cm nylon filters containing PCR products corresponding to 4290 ORFs from *E. coli*. 10 ngs of each are spotted every 1.5 mm on the filter. Single stranded labeled cDNAs are prepared for hybridization to the array (no second strand synthesis or amplification step is done) and placed in contact with the filter. Thus the labeled cDNAs are of "antisense" orientation. Quantitative analysis is done by phosphorimager.

Hybridization of cDNA made from a sample of total cell mRNA to such an array followed by detection of binding by one or more of various techniques known to those in the art results in a signal at each location on the array to which cDNA hybridized. The intensity of the hybridization signal obtained at each location in the array thus reflects the amount of mRNA for that specific gene that was present in the sample. Comparing the results obtained for mRNA isolated from cells grown under different conditions thus allows for a comparison of the relative amount of expression of each individual gene during growth under the different conditions.

Gene expression arrays may be used to analyze the total mRNA expression pattern at various time points after induction of an antisense nucleic acid complementary to a proliferation-required gene. Analysis of the expression pattern indicated by hybridization to the array provides

information on whether or not the target gene of the antisense nucleic acid is being affected by antisense induction, how quickly the antisense is affecting the target gene, and for later timepoints, what other genes are affected by antisense expression. For example, if the antisense is directed against a gene for ribosomal protein L7/L12 in the 50S subunit, its targeted mRNA may disappear first and then other mRNAs may be observed to increase, decrease or stay the same. Similarly, if the antisense is directed against a different 50S subunit ribosomal protein mRNA (e.g. L25), that mRNA may disappear first followed by changes in mRNA expression that are similar to those seen with the L7/L12 antisense expression. Thus, the mRNA expression pattern observed with an antinsense nucleic acid complementary to a proliferation required gene may identify other proliferation-required nucleic acids in the same pathway as the target of the antisense nucleic acid. In addition, the mRNA expression patterns observed with candidate drug compounds may be compared to those observed with antisense nucleic acids against a proliferation-required nucleic acid. If the mRNA expression pattern observed with the candidate drug compound is similar to that observed with the antisense nucleic acid, the drug compound may be a promising therapeutic candidate. Thus, the assay would be useful in assisting in the selection of candidate drug compounds for use in screening methods such as those described below.

In cases where the source of nucleic acid deposited on the array and the source of the nucleic acid being hybridized to the array are from two different organisms, gene expression arrays can identify homologous genes in the two organisms.

The present invention also contemplates additional methods for screening other microorganisms for proliferation-required genes. In this embodiment, the conserved portions of sequences identified as proliferation-required can be used to generate degenerate primers for use in the polymerase chain reaction (PCR). The PCR technique is well known in the art. The successful production of a PCR product using degenerate probes generated from the sequences identified herein would indicate the presence of a homologous gene sequence in the species being screened. This homologous gene is then isolated, expressed, and used as a target for candidate antibiotic compounds. In another aspect of this embodiment, the homologous gene is expressed in an autologous organism or in a heterologous organism in such a way as to alter the level or activity of a homologous gene required for proliferation in the autologous or heterologus organism. In still another aspect of this embodiment, the homologous gene or portion is expressed in an antisense orientation in such a way as to alter the level or activity of a nucleic acid required for proliferation of an autologous or heterologous organism.

The homologous sequences to proliferation-required genes identified using the techniques described herein may be used to identify proliferation-required genes of organisms other than *E. coli*, to inhibit the proliferation of organisms other than *E. coli* by inhibiting the activity or reducing the amount of the identified homologous nucleic acid or polypeptide in the organism other than *E. coli*, or to identify compounds which inhibit the growth of organisms other than *E. coli* as described below.

In another embodiment of the present invention, *E. coli* sequences identified as required for proliferation are transferred to expression vectors capable of function within non-*E coli* species. As would be appreciated by one of ordinary skill in the art, expression vectors must contain certain elements that are species specific. These elements can include promoter sequences, operator sequences, repressor genes, origins of replication,selectable marker genes, ribosomal binding sequences, termination sequences, and others. To use the identified exogenous sequences of the present invention, one of ordinary skill in the art would know to use standard molecular biology techniques to isolate vectors containing the sequences of interest from cultured bacterial cells, isolate and purify those sequences, and subclone those sequences into an expression vector adapted for use in the species of bacteria to be screened.

Expression vectors for a variety of other species are known in the art. For example, Cao et al. report the expression of steroid receptor fragments in *Staphylococcus aureus*. J. Steroid Biochem Mol Biol. 44(1):1-11 (1993). Also, Pla et al. have reported an expression vector that is functional in a number of relevant hosts including: *Salmonella typhimurium, Pseudomonas putida, and Pseudomonas aeruginosa*. J. Bacteriol. 172(8):4448-55 (1990). These examples demonstrate the existence of molecular biology techniques capable of constructing expression vectors for the species of bacteria of interest to the present invention.

Following the subcloning of the identified nucleic acid sequences into an expression vector functional in the microorganism of interest, the identified nucleic acid sequences are conditionally transcribed to assay for bacterial growth inhibition. Those expression vectors found to contain sequences that, when transcribed, inhibit bacterial growth are compared to the known genomic sequence of the pathogenic microorganism being screened or, if the homologous sequence from the organism being screened is not known, it may be identified and isolated by hybridization to the proliferation-required *E. coli* sequence interest or by amplification using primers based on the proliferation-required *E. coli* sequence of interest as described above.

The antisense sequences from the second organism which are identified as described above may then be operably linked to a promoter, such as an inducible promoter, and introduced into the second organism. The techniques described herein for identifying *E. coli* genes required for proliferation may thus be employed to determine whether the identified sequences from a second organism inhibit the proliferation of the second organism.

Antisense nucleic acids required for the proliferation of organisms other than *E. coli* or the genes corresponding thereto, may also be hybridized to a microarray containing the *E. coli* ORFs to gauge the homology between the *E. coli* sequences and the proliferation-required nucleic acids from other organisms. For example, the proliferation-required nucleic acid may be from *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis),*

Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis or any species falling within the genera of any of the above species. The proliferation-required nucleic acids from an organism other than *E. coli* may be hybridized to the array under a variety of conditions which permit hybridization to occur when the probe has different levels of homology to the sequence on the microarray. This would provide an indication of homology across the organisms as well as clues to other possible essential genes in these organisms.

In still another embodiment, the exogenous nucleic acid sequences of the present invention that inhibit bacterial growth or proliferation can be used as antisense therapeutics for killing bacteria. The antisense sequences can be complementary to the proliferation-required genes whose sequence corresponds to the exogenous nucleic acid probes identified here (i.e. the antisense nucleic acid may hybridize to the gene or a portion thereof). Alternatively, antisense therapeutics can be complementary to operons in which proliferation-required genes reside (i.e. the antisense nucleic acid may hybridize to any gene in the operon in which the proliferation-required genes reside). Further, antisense therapeutics can be complementary to a proliferation-required gene or portion thereof with or without adjacent noncoding sequences, an intragenic sequence (i.e. a sequence within a gene), an intergenic sequence (i.e. a sequence between genes), a sequence spanning at least a portion of two or more genes, a 5' noncoding region or a 3' noncoding region located upstream or downstream from the actual sequence that is required for bacterial proliferation or an operon containing a proliferation-required gene.

In addition to therapeutic applications, the present invention encompasses the use of nucleic acid sequences complementary to sequences required for proliferation as diagnostic tools. For example, nucleic acid probes complementary to proliferation-required sequences that are specific for particular species of microorganisms can be used as probes to identify particular microorganism species in clinical specimens. This utility provides a rapid and dependable method by which to identify the causative agent or agents of a bacterial infection. This utility would provide clinicians the ability to prescribe species specific antimicrobial compounds to treat such infections. In an extension of this utility, antibodies generated against proteins translated from mRNA transcribed from proliferation-required sequences can also be used to screen for specific microorganisms that produce such proteins in a species-specific manner.

The following examples teach the genes of the present invention and a subset of uses for the *E. coli* genes identified as required for proliferation. These examples are illustrative only and are not intended to limit the scope of the present invention.

EXAMPLES

The following examples are directed to the identification and exploitation of *E. coli* genes required for proliferation. Methods of gene identification are discussed as well as a variety of methods to utilize the identified sequences.

Genes Identified as Required for Proliferation of *E. coli*

Exogenous nucleic acid sequences were cloned into an inducible expression vector and assayed for growth inhibition activity. Example 1 describes the examination of a library of exogenous nucleic acid sequences cloned into the IPTG-inducible expression vector pLEX5BA (Krause et al., J. Mol. Biol. 274: 365 (1997)) or a modified version of pLEX5BA, pLEX5BA-3' in which a synthetic linker containing a T7 terminator was ligated between the PstI and HindIII sites of pLEX5BA. In particular, to construct pLEX5BA-3', the following oligonucleotides were annealed and inserted into the PstI and HindIII sites of pLEX5BA:

```
5'-GTCTAGCATAACCCCTTGGGGCCTCTAAACGGGTCCTTGAGGGGTTTTTTGA-3' (SEQ ID
NO: 480)

5'-AGCTTCAAAAAACCCCTCAAGGACCCGTTTAGAGGCCCCAAGGGGTTAT
GCTAGACTGCA-3' (SEQ ID NO: 481)
```

Random fragments of *E. coli* genomic DNA were generated by DNAseI digestion or sonication, filled in with T4 polymerase, and cloned into the SmaI site of pLEX5BA or pLEX5BA-3'. Upon activation or induction, the expression vectors produced an RNA molecule corresponding to the subcloned exogenous nucleic acid sequences. The RNA product was in an antisense orientation with respect to the *E. coli* genes from which it was originally derived. This antisense RNA then interacted with sense mRNA produced from various *E. coli* genes and interfered with or inhibited the translation of the sense messenger RNA (mRNA) thus preventing protein production from these sense mRNA molecules. In cases where the sense mRNA encoded a protein required for the proliferation, bacterial cells containing an activated expression vector failed to grow or grew at a substantially reduced rate. Similar results have also been obtained in cases where the gene encodes a non-translated RNA, such as a ribosomal RNA.

It will be appreciated that vectors other than pLEX5BA or pLEX5BA-3' may be used to transcribe the genomic DNA inserts. In addition, it will be appreciated that, if desired, pLEX5BA or pLEX5BA-3' may be modified to introduce features such as stop codons in all three reading frames downstream of the genomic DNA inserts to ensure that if the genomic DNA insert encodes a polypeptide (i.e. the insert is in the sense orientation rather than the antisense orientation or the

insert is in the antisense orientation but contains a cryptic ORF) translation of the polypeptide will terminate shortly after the genomic insert.

EXAMPLE 1

Inhibition of Bacterial Proliferation after IPTG induction

To study the effects of transcriptional induction in liquid medium, growth curves were carried out by back diluting cultures 1:200 into fresh media with or without 1 mM IPTG and measuring the OD_{450} every 30 minutes (min). To study the effects of transcriptional induction on solid medium, 10^2, 10^3, 10^4, 10^5, 10^6, 10^7 and 10^8 fold dilutions of overnight cultures were prepared. Aliquots of from 0.5 to 3 µl of these dilutions were spotted on selective agar plates with or without 1 mM IPTG. After overnight incubation, the plates were compared to assess the sensitivity of the clones to IPTG.

Of the numerous clones tested, some clones were identified as containing a sequence that inhibited *E. coli* growth after IPTG induction. Accordingly, the gene to which the inserted nucleic acid sequence corresponds, or a gene within the operon containing the inserted nucleic acid, may be required for proliferation in *E. coli*.

Characterization of Isolated Clones Negatively Affecting *E. coli* Proliferation

Following the identification of those inserts that, upon expression, negatively impacted *E. coli* growth or proliferation, the inserts were isolated and subjected to nucleic acid sequence determination.

EXAMPLE 2

Nucleic Acid Sequence Determination of Identified Clones Expressing Nucleic Acid Fragments with Detrimental Effects of *E. coli* Proliferation

The nucleotide sequences for the exogenous identified sequences were determined using plasmid DNA isolated using QIAPREP (Qiagen, Valencia, CA) and methods supplied by the manufacturer. The primers used for sequencing the inserts were 5' - TGTTTATCAGACCGCTT - 3' (SEQ ID NO: 1) and 5' - ACAATTTCACACAGCCTC - 3' (SEQ ID NO: 2). These sequences flank the polylinker in pLEX5BA. Sequence identification numbers (SEQ ID NOs) for the identified inserts are listed in Table I and discussed below.

EXAMPLE 3

Comparison Of Isolated Sequences to Known Sequences

The nucleic acid sequences of the subcloned fragments obtained from the expression vectors discussed above were compared to known *E. coli* sequences in GenBank using BLAST version 1.4 or version 2.0.6 using the following default parameters: Filtering off, cost to open a gap=5, cost to extend a gap=2, penalty for a mismatch in the blast portion of run=-3, reward for a match in the blast portion of run=1, expectation value (e)=10.0, word size=11, number of one-line descriptions=100, number of alignments to show (B)=100. BLAST is described in Altschul, J Mol Biol. 215:403-10 (1990). Expression vectors were found to contain nucleic acid sequences in both the sense and antisense orientations. The presence of known genes, open reading frames, and ribosome binding sites was determined by comparison to public databases holding genetic information and various computer

programs such as the Genetics Computer Group programs FRAMES and CODONPREFERENCE. Clones were designated as "antisense" if the cloned fragment was oriented to the promoter such that the RNA transcript produced was complementary to the expressed mRNA (or non-translated RNA) from a chromosomal locus. Clones were designated as "sense" if they coded for an RNA fragment that was identical to a portion of a wild type mRNA from a chromosomal locus.

The sequences described in Examples 1-2 that inhibited bacterial proliferation and contained gene fragments in an antisense orientation are listed in Table I. This table lists each identified sequence by: a sequence identification number; a Molecule Number; a gene to which the identified sequence corresponds, listed according to the National Center for Biotechnology Information (NCBI), Blattner (Science 277:1453-1474(1997); also contains the *E. coli* K-12 genome sequence), or Rudd (Micro. and Mol. Rev. 62:985-1019 (1998)), nomenclatures. The CONTIG numbers for each identified sequence is shown, as well as the location of the first and last base pairs located on the *E. coli* chromosome. A Molecule Number with a "**" indicates a clone corresponding to an intergenic sequence.

TABLE I

Identified Clones with Corresponding Genes and Operons

SeqID	Molecule Number	Clone name	Contig	start	stop	Gene name (Blattner)	Gene name (NCBI)
1	EcXA118a	E1M10000131C06	AE000408	5299	5440	b3310	rplN
2	EcXA118b	E1M10000152F04	AE000408	5300	5452	b3310	rplN
3	EcXA118c	E1M10000152H04	AE000408	5311	5450	b3310	rplN
4	EcXA118d	E1M10000153H03	AE000408	5299	5475	b3310	rplN
5	EcXA119	E1M10000129F10	AE000372	2407	2153	b2883	b2883
6	EcXA120	E1M10000129G04	AE000248	2494	2888	b1509	b1509
7	EcXA121	E1M10000119D03	AE000491	1722	2001	b4191	yjfQ
8	EcXA122a	1029-M5	AE000300	232	1	b2108	yehA
			AE000299	10036	9848	b2107	b2107
9	EcXA122b	X3S208-17	AE000300	186	1	b2108	yehA
			AE000299	10036	9870	b2107	b2107
10	EcXA122c	E1M10000124C07	AE000299	10036	9896	b2107	b2107
			AE000300	78	1	b2108	yehA
11	EcXA122d	E1M10000125D06	AE000300	203	1	b2108	yehA
			AE000299	10036	9892	b2107	b2107
12	EcXA123	E1M10000106A08	AE000410	7832	7968	b3343	yheL
13	EcXA124	E1M10000106D02	AE000408	5951	6225	b3311	rpsQ
			AE000408	5951	6225	b3312	rpmC
14	EcXA125	E1M10000124D09	AE000465	7071	7309	b3901	yiiL
15	EcXA126	E1M10000124E02	AE000441	9429	9316	b3644	yicC
16	EcXA127a	E1M10000124E06	AE000369	1145	856	b2851	b2851
17	EcXA127b	E1M10000159A11	AE000369	1145	763	b2851	b2851
18	EcXA128	E1M10000124G04	AE000171	6676	6427	b0679	nagE
19	EcXA129	E1M10000127C09	AE000123	1270	1506	b0135, b0136	yadC, yadK
20	EcXA130	E1M10000146E05	AE000450	9592	9828	b3737, b3738	atpE, atpB
21	EcXA131	E1M10000162C01	AE000383	11211	11350	b3016	b3016

SeqID	Molecule Number	Clone name	Contig	start	stop	Gene name (Blattner)	Gene name (NCBI)
22	EcXA132	E1M10000163A04	AE000243	5440	5650	b1465	narV
23	EcXA133a	E1M10000120F06	AE000408	5273	5425	b3309, b3310	rplX, rplN
24	EcXA133b	E1M10000132H04	AE000408	5263	5422	b3309, b3310	rplX, rplN
25	EcXA133c	E1M10000172C05	AE000408	5190	5433	b3309, b3310	rplX, rplN
26	EcXA134	E1M10000162H06	AE000402	4689	4795	b3229	sspA
27	EcXA136a	E1M10000147B03	AE000373	1	251	b2892	recJ
			AE000372	11967	12144	b2891	prfB
28	EcXA136b	E1M10000155F04	AE000373	1	372	b2892	recJ
			AE000372	12002	12144	b2891	prfB
29	EcXA137	E1M10000142F12	AE000132	9172	8451	b0243, b0244	proA, thrW
30	EcXA138	E1M10000148B09	AE000393	10354	10051	b3121	yhaC
31	EcXA139	E1M10000143A12	AE000189	3600	4024	b0874	b0874
32	EcXA139b	E1M10000169H02	AE000189	3646	4026	b0874	b0874
33	EcXA140a	E1M10000143C10	AE000445	1	260	b3672	b3672
			AE000444	12573	12697	b3671	ilvB
34	EcXA140b	E1M10000151F10	AE000445	1	378	b3672	b3672
			AE000444	12576	12697	b3671	ilvB
35	EcXA141	E1M10000143G09	AE000294	9227	9498	b2035	rfc
36	EcXA142	E1M10000159C10	AE000408	1	199	b3299	rpmJ
			AE000407	10488	10601	b3298	rpsM
37	EcXA143	E1M10000159F03	AE000368	7705	7317	b2846	b2846
38	EcXA144a	E1M10000160A03	AE000257	545	368	b1613	manA
39	EcXA144b	E1M10000132F02	AE000257	585	395	b1613	manA
40	EcXA144c	E1M10000153E04	AE000257	552	395	b1613	manA
41	EcXA145	E1M10000160A06	AE000366	5350	5533	b2827	thyA
42	EcXA146	E1M10000160C02	AE000449	9694	9815	b3725	pstB
43	EcXA147	E1M10000160C03	AE000211	6964	7126	b1111	b1111
44	EcXA148	E1M10000160E10	AE000455	10737	10577	b3794	rffM
45	EcXA149a	E1M10000160F11	AE000374	7930	7680	b2912	ygfA
46	EcXA149b	E1M10000153F05	AE000374	7910	7625	b2912	ygfA
47	EcXA150	E1M10000160H03	AE000442	5556	5377	b3650	spoT
48	EcXA151	E1M10000144E03	AE000429	8099	8352	b3528	dctA
49	EcXA152	E1M10000144G07	AE000308	1086	1385	b2182	bcr
50	EcXA153	E1M10000144B01	AE000477	4261	4055	b4038	yjbI
51	EcXA154	E1M10000150E02	AE000290	5618	5406	b1983	b1983
52	EcXA155	E1M10000150E06	AE000339	116	496	b2522	sseB
53	EcXA156	E1M10000156B08	AE000463	2852	2699	b3874	yihN
54	EcXA157	E1M10000156D07	AE000408	7274	7464	b3314, b3315	rpsC, rplV
55	EcXA158	E1M10000156G12	AE000338	5181	5370	b2519	b2519
56	EcXA159	E1M10000151G10	AE000154	7217	7618	b0482	b0482
57	EcXA160	E1M10000166G06	AE000269	4488	5074	b1744	b1744
58	EcXA161	869.A23	AE000146	3702	3950	b0398	sbcD
59	EcXA162	E1M10000112F05	AE000497	7921	8156	b4267	yjgV
60	EcXA163	E1M10000118B05	AE000447	8832	8741	b3705	yidC
61	EcXA164	E1M10000118C04	AE000314	6749	6688	b2243	glpC
62	EcXA165	E1M10000118C05	AE000170	219	290	b0655	ybeJ
63	EcXA166	E1M10000118G06	AE000461	5835	5770	b3859	yihE
64	EcXA167	E1M10000119A05	AE000298	9388	9521	b2093	gatB
65	EcXA168	E1M10000123E09	AE000267	4622	4413	b1721	b1721
66	EcXA169	E1M10000123F11	AE000408	2650	2767	b3303	rpsE
67	EcXA170	E1M10000129G01	AE000143	3523	3913	b0363	b0363
68	EcXA171	E1M10000132G08	AE000189	6799	6699	b0876	b0876
69	EcXA172	E1M10000139B11	AE000323	3697	3531	b2345	b2345
70	EcXA173	E1M10000153A09	AE000424	6886	6968	b3485	yhhJ

SeqID	Molecule Number	Clone name	Contig	start	stop	Gene name (Blattner)	Gene name (NCBI)
71	EcXA174	E1M10000153C04	AE000377	8932	8832	b2945	endA
72	EcXA175	E1M10000154F03	AE000279	7957	7837	b1854	pykA
73	EcXA176	E1M10000157B02	AE000485	4932	5094	b4125	yjdH
74	EcXA177	E1M10000160H05	AE000326	3140	3035	b2378	b2378
75	EcXA178	E1M10000163H01	AE000113	1	266	b0023	rpsT
76	EcXA179	E1M10000164G04	AE000262	233	1	b1665	valV
77	EcXA180	E1M10000168F02	AE000480	3110	3350	b4066	yjcF
78	EcXA181	E1M10000169D08	AE000469	10535	10604	b3956	ppc
79	EcXA182	E1M10000169H07	AE000376	10447	10586	b2935	tktA
80	EcXA183	E1M10000169H08	AE000118	8416	8228	b0084	ftsI
81	EcXA184	E1M10000172C01	AE000358	8498	8844	b2750	cysC
82	EcXA185	E1M10000176C01	AE000407	10053	10119	b3297	rpsK
83	EcXA186	E1M10000176F01	AE000135	987	812	b0270	yagG
84	EcXA187	P319-3.M16	AE000317	2279	1872	b2271	b2271
85	EcXA188	P347.2	AE000211	7845	8005	b1113	b1113
86	EcXA189a	E1M10000168A04	AE000277	2048	1861	b1828	b1828
87	EcXA189b	E1M10000125F09	AE000277	2042	1868	b1828	b1828
88	EcXA189c	E1M10000113F02	AE000277	2050	1857	b1828	b1828
89	EcXA190a	E1M10000161D04	AE000387	49	312	b3052	b3052
90	EcXA190b	E1M10000155B11	AE000387	49	285	b3052	b3052
91	EcXA191a	E1M10000167H09	AE000451	18	220	b3740	gidB
92	EcXA191b	E1M10000120H05	AE000451	1	188	b3740	gidB
93	EcXA192	E1M10000154H07	AE000452	3326	3106	b3750, b3751	rbsC, rpsB

EXAMPLE 4

Identification of Genes and their Corresponding Operons Affected by Antisense Inhibition

The sequencing of the entire *E. coli* genome is described in Blattner et al., Science 277:1453-1474(1997) and the sequence of the genome is listed in GenBank Accession No.U00096. The operons to which the proliferation-inhibiting nucleic acids correspond were identified using RegulonDB and information in the literature. The coordinates of the boundaries of these operons on the *E. coli* genome are listed in Table III. Table II lists the molecule numbers of the inserts containing the growth inhibiting nucleic acid fragments, the genes in the operons corresponding to the inserts, the SEQ ID NOs of the genes containing the inserts, the SEQ ID NOs of the proteins encoded by the genes, the start and stop points of the genes on the *E. coli* genome, the orientation of the genes on the genome, whether the operons are predicted or documented, and the predicted functions of the genes. The identified operons, their putative functions, and whether or not the genes are presently thought to be required for proliferation are discussed below.

Functions for the identified genes were determined by using either Blattner functional class designations or by comparing identified sequence with known sequences in various databases. A variety of biological functions were noted for the genes to which the clones of the present invention correspond. The functions for the genes of interest appear in Table II.

The proteins that are listed in Table II are involved in a wide range of biological functions.

TABLE II

All Operon Data with Whole Chromosome Coordinates

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
EcXA118a-d	*rpmJ*	94	287	3440255	3440371	<	documented	50S ribosomal subunit protein X (L36)
	prlA	95	288	3440403	3441734	<		*SecY*: multispanning membrane protein, translocator of proteins
	rplO	96	289	3441742	3442176	<		50S ribosomal subunit protein L15
	rpmD	97	290	3442180	3442359	<		50S ribosomal subunit protein L30
	rpsE	98	291	3442363	3442866	<		eps, spc, spcA; 30S ribosomal subunit protein S5
	rplR	99	292	3442881	3443234	<		50S ribosomal subunit protein L18
	rplF	100	293	3443244	3443777	<		50S ribosomal subunit protein L6; gentamicin sensitivity
	rpsH	101	294	3443790	3444182	<		30S ribosomal subunit protein S8
	rpsN	102	295	3444216	3444521	<		30S ribosomal subunit protein S14
	rplE	103	296	3444536	3445075	<		50S ribosomal subunit protein L5
	rplX	104	297	3445090	3445404	<		50S ribosomal subunit protein L24
	rplN	105	298	3445415	3445786	<		50S ribosomal subunit protein L14
EcXA119	*b2882*	106	299	3022315	3023772	>	predicted	Unknown
	b2883	107	300	3023787	3025106	>		Unknown
	b2884	108	301	3025142	3025711	>		Unknown
	b2885	109	302	3025678	3026508	>		Unknown
EcXA120	*b1509*	110	303	1590689	1592089	<	predicted	Unknown
	ydeK	111	304	1592133	1596110	<		Unknown
EcXA121	*yjfQ*	112	305	4415276	4416031	<	predicted	Unknown
EcXA122	*b2106*	113	306	2183937	2184761	>	predicted	Unknown
	b2107	114	307	2184800	2185318	>		Unknown
	yehA	115	308	2185400	2186434	<	predicted	Unknown
	yehB	116	309	2186450	2188930	<		Unknown

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
	yehC	117	310	2188946	2189665	<		Unknown
	yehD	118	311	2189700	2190242	<		Unknown
EcXA123	*yheL*	119	312	3472315	3472602	<	predicted	Unknown
	b3344	120	313	3472610	3472969	<		Unknown
	yheN	121	314	3472969	3473355	<		Unknown
	b3346	122	315	3473355	3474089	<		Unknown
EcXA124	*rpsQ*	123	316	3445951	3446205	<	documented	neaA; 30S ribosomal subunit protein S17
	rpmC	124	317	3446205	3446396	<		50S ribosomal subunit protein L29
	rplP	125	318	3446396	3446806	<		50S ribosomal subunit protein L16
	rpsC	126	319	3446819	3447520	<		30S ribosomal subunit protein S3
	rplV	127	320	3447538	3447870	<		eryB; erythromycin sensitivity; 50S ribosomal subunit protein L22
	rpsS	128	321	3447885	3448163	<		30S ribosomal subunit protein S19
	rplB	129	322	3448180	3449001	<		50S ribosomal subunit protein L2
	rplW	130	323	3449019	3449321	<		50S ribosomal subunit protein L23
	rplD	131	324	3449318	3449923	<		eryA; 50S ribosomal subunit protein L4; erythromycin sensitivity
	rplC	132	325	3449934	3450563	<		50S ribosomal subunit protein L3
	rpsJ	133	326	3450596	3450907	<		nusE; 30S ribosomal subunit protein S10
EcXA125	*yiiL*	134	327	4090705	4091019	<	predicted	Unknown
EcXA126	*yicC*	135	328	3814303	3815166	>	predicted	Unknown
EcXA127a-b	*b2851*	136	329	2989290	2989781	>	predicted	Unknown
EcXA128	*nagE*	137	330	703167	705113	>	documented	pstN; N-acetylglucosamine-specific enzyme II of phosphotransferase system
EcXA129	*yadC*	138	331	149715	150953	<	predicted	Unknown
	yadK	139	332	151003	151599	<		Unknown
	yadL	140	333	151626	152231	<		Unknown
	yadM	141	334	152243	152854	<		Unknown

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
EcXA130	*atpC*	142	335	3913181	3913600	<	documented	papG, uncC; membrane-bound ATP synthase, F1 sector, -subunit (EC 3.6.1.3)
	atpD	143	336	3913621	3915003	<		papB, uncD; membrane-bound ATP synthase, F1 sector, -subunit (EC 3.6.1.3)
	atpG	144	337	3915030	3915893	<		papC, uncG; membrane-bound ATP synthase, F1 sector, -subunit (EC 3.6.1.3)
	atpA	145	338	3915944	3917485	<		papA, uncA; membrane-bound ATP synthase, F1 sector, -subunit (EC 3.6.1.3)
	atpH	146	339	3917498	3918031	<		papE, uncH; membrane-bound ATP synthase, F1 sector, -subunit (EC 3.6.1.3)
	atpF	147	340	3918046	3918516	<		papF, uncF; membrane-bound ATP synthase, F0 sector, subunit b (EC 3.6.1.3)
	atpE	148	341	3918578	3918817	<		papH, uncE; membrane-bound ATP synthase, F0 sector, subunit c; DCCD (EC 3.6.1.3)
	atpB	149	342	3918864	3919679	<		papD, uncB; membrane-bound ATP synthase, F0 sector, subunit a (EC 3.6.1.3)
	atpI	150	343	3919688	3920080	<		uncI; membrane-bound ATP synthase subunit, F1-F0-type proton-ATPase (EC 3.6.1.34)
EcXA131	*b3015*	151	344	3156944	3158185	<	predicted	Unknown
	b3016	152	345	3158185	3159162	<		Unknown
EcXA132	*narV*	153	346	1533961	1534641	<	documented	regulationCryptic nitrate reductase II, -subunit
	narW	154	347	1534638	1535333	<		regulationCryptic nitrate reductase II, -subunit
	narY	155	348	1535333	1536877	<		regulationCryptic NR II, -subunit
	narZ	156	349	1536874	1540614	<		regulationCryptic NR II, -subunit
EcXA133a-c							Same operon as EcXA118	

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
EcXA134	*sspB*	157	350	3373914	3374411	<	documented	Stress response protein
	sspA	158	351	3374417	3375055	<		pog; stress response protein
EcXA136a-b	*lysS*	159	352	3031677	3033194	<	documented	asuD herC; lysyl tRNA synthetase, constitutive
	prfB	160	353	3033204	3034302	<		supK; peptide chain release factor 2
	recJ	161	354	3034393	3036126	<	predicted	Single-stranded DNA-specific exonuclease, 5 -3
	dsbC	162	355	3036132	3036842	<		xprA; periplasmic disulfide oxidoreductase, protein disulfide isomerase
	xerD	163	356	3036867	3037763	<		xprB; recombinase, site specific
EcXA137	*proB*	164	357	259612	260715	>	documented	pro2; -glutamyl kinase (EC 2.7.2.11)
	proA	165	358	260727	261980	>		pro1; -glutamyl phosphate reductase (EC 1.2.1.41)
EcXA138	*yhaB*	166	359	3265474	3266034	>	predicted	
	yhaC	167	360	3266056	3267243	>		
EcXA139a-b	*b0874*	168	361	913181	914128	<	predicted	
EcXA140a-b	*ilvN*	169	362	3848429	3848719	<	documented	Acetohydroxy acid synthase I, small subunit (EC 4.1.3.18); valine sensitive small subunit;positive regulator for thr and ilv
	ilvB	170	363	3848723	3850411	<		Acetolactate synthase I, valine sensitive (EC 4.1.3.18)
	b3672	171	364	3850517	3850615	<		
EcXA141	*yi52_7*	172	365	2099917	2100933	<	predicted	nucleotide sequence and phylogenetic relationships of a new E. coli insertion element. Schwartz E
	yefJ	173	366	2100938	2101411	<		Unknown
	yefI	174	367	2101413	2102531	<		Unknown
	yefH	175	368	2102516	2103106	<		Unknown

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
	yefG	176	369	2103087	2104079	<		Unknown
	rfc	177	370	2104082	2105248	<		O-antigen polymerase
	yefE	178	371	2105248	2106351	<		Unknown
EcXA142							Same operon as EcXA118	
EcXA143	*b2846*	179	372	2985498	2986190	>	predicted	Unknown
EcXA144a-c	*manA*	180	373	1686600	1687775	>	predicted	pmi; mannosephosphate isomerase (EC 5.3.1.8)
EcXA145	*thyA*	181	374	2962383	2963177	<	documented	Aminopterin, trimethoprim resistance; thymidylate synthetase (EC 2.1.1.45)
	lgt	182	375	2963184	2964059	<		transferaseumpA; phosphatidylglycerol:prolipoprotein diacylglycerol transferase
EcXA146	*phoU*	183	376	3904481	3905206	<	documented	phoT; P uptake, high-affinity P-specific transport system, regulatory gene
	pstB	184	377	3905221	3905994	<		phoT; high-affinity P-specific transport; cytoplasmic ATP-binding protein
	pstA	185	378	3906177	3907067	<		R2pho, phoR2b, phoT; high-affinity P-specific transport
	pstC	186	379	3907067	3908026	<		phoW; high-affinity P-specific transport; cytoplasmic membrane component
	pstS	187	380	3908113	3909153	<		phoR2, nmpA, phoR2a, phoS, R2pho; high-affinity P-specific transport; periplasmic Pbinding
EcXA147	*b1111*	188	381	1167423	1168133	<	predicted	Unknown
EcXA148	*yifH*	189	382	3972208	3972753	>	predicted	Unknown
	yifI	190	383	3972758	3973888	>		Unknown
	yifJ	191	384	3973890	3975140	>		Unknown
	rffT	192	385	3976214	3977566	>		Synthesis of enterobacterial common antigen; Fuc4NAc transferase

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
	rffM	193	386	3977569	3978309	>		UDP-ManNAcA transferase
	b4404	194	387	3975137	3975361	>		Unknown
	b4405	195	388	3975603	3976217	>		Unknown
EcXA149a-b	*ygfA*	196	389	3054261	3054809	>	predicted	Unknown
EcXA150	*rpoZ*	197	390	3819733	3820008	>	documented	spoS; RNA polymerase,
	spoT	198	391	3820027	3822135	>		Guanosine 5 -diphosphate, 3 -diphosphate pyrophosphatase, ppGpp synthetase II activity
	spoU	199	392	3822142	3822831	>		
	recG	200	393	3822837	3824918	>		spoV? branch migration of Holliday junctions, junction-specific DNA helicase (see ruvABC)
EcXA151	*dctA*	201	394	3679791	3681077	<	predicted	Uptake of C-4 dicarboxylic acids; 3-fluoromalate resistance, D-tartrate resistant
EcXA152	*bcr*	202	395	2276590	2277780	<	predicted	bicA, bicR, sur, suxA; transmembrane; affects sulfathiazole-sulfonamide resistance
	yejD	203	396	2277808	2278503	<		Unknown
EcXA153	*yjbI*	204	397	4248534	4249862	>	predicted	Unknown
EcXA154	*b1983*	205	398	2054880	2055596	>	predicted	Unknown
EcXA155	*sseB*	206	399	2652177	2652962	<	predicted	Enhances serine sensitivity (inhibits homoserine deHase) on lactate; weaker than sseA
EcXA156	*yihN*	207	400	4059826	4061091	>	predicted	Unknown
EcXA157							Same operon as EcXA124	Unknown
EcXA158	*b2519*	208	401	2643033	2645345	<	predicted	Unknown
	b2520	209	402	2645346	2650307	<		Unknown
EcXA159	*b0482*	210	403	506510	507304	<	predicted	Unknown

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
EcXA160	*b1744*	211	404	1823979	1824947	<	predicted	Unknown
	b1745	212	405	1824940	1826283	<		Unknown
	b1746	213	406	1826280	1827758	<		Unknown
	b1747	214	407	1827755	1828789	<		Unknown
	b1748	215	408	1828786	1830006	<		Unknown
EcXA161	*sbcC*	216	409	411831	414977	<	predicted	Cosuppressor with sbcB of recB recC mutations
	sbcD	217	410	414974	416176	<		Cosuppressor with sbcB of recB recC mutations
EcXA162	*yjgU*	218	411	4490155	4490919	<	predicted	Unknown
	yjgV	219	412	4490943	4491974	<		Unknown
EcXA163	*yidC*	220	413	3882705	3884351	>	predicted	Unknown
EcXA164	*glpA*	221	414	2350667	2352295	>	documented	Glycerol-3-phosphate dehydrogenase (anaerobic) large subunit (EC 1.1.99.5)
	glpB	222	415	2352285	2353544	>		sn-Glycerol-3-phosphate dehydrogenase (anaerobic) subunit (EC 1.1.99.5); membraneanchor
	glpC	223	416	2353541	2354731	>		sn-Glycerol-3-phosphate dehydrogenase (anaerobic) small subunit (EC 1.1.99.5)
EcXA165	*ybeJ*	224	417	686062	686970	<	predicted	Unknown
EcXA166	*yihE*	225	418	4039996	4040982	>	predicted	Unknown
	dsbA	226	419	4040999	4041625	>		iarA, ppfA; disulfide oxidoreductase, periplasmic protein disulfide-isomerase; role incytochrome c synthesis (EC 5.3.4.1)
EcXA167	*gatR_2*	227	420	2169417	2169755	<	documented	
	gatD	228	421	2169855	2170895	<		Galactitol-1-phosphate dehydrogenase
	gatC	229	422	2170943	2172298	<		Galactitol-specific enzyme IIC of PTS
	gatB	230	423	2172302	2172586	<		Galactitol-specific enzyme IIB of PTS

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
	gatA	231	424	2172617	2173069	<		Galactitol-specific enzyme IIA of phosphotransferase system (PTS)
	gatZ	232	425	2173079	2174341	<		Function unknown
	gatY	233	426	2174370	2175230	<		D-Tagatose-1,6-bisphosphate aldolase
EcXA168	*b1720*	234	427	1801118	1801591	>	predicted	Unknown
	b1721	235	428	1801602	1803017	>		Unknown
EcXA169							Same operon as EcXA118	Unknown
EcXA170	*b0362*	236	429	381728	382114	<	predicted	Unknown
	b0363	237	430	381963	383159	<		Unknown
EcXA171	*b0876*	238	431	915696	917354	>	predicted	Unknown
EcXA172	*b2345*	239	432	2461032	2462090	>	predicted	Unknown
EcXA173	*yhhJ*	240	433	3623310	3624437	<	predicted	Unknown
	b3486	241	434	3624434	3627118	<		Unknown
	yhiI	242	435	3627165	3628232	<		Unknown
EcXA174	*endA*	243	436	3088366	3089073	>	predicted	DNA-specific endonuclease I; extensive DNA breakdown
EcXA175	*pykA*	244	437	1935673	1937115	>	predicted	Pyruvate kinase A (II); (EC 2.7.1.40)
EcXA176	*yjdG*	245	438	4346893	4347612	<	predicted	Unknown
	yjdH	246	439	4347609	4349240	<		Unknown
EcXA177	*b2378*	247	440	2493599	2494585	>	predicted	Unknown
EcXA178	*rpsT*	248	441	20815	21078	<	predicted	sups20; 30S ribosomal subunit protein S20
EcXA180	*yjcF*	249	442	4279362	4280654	<	predicted	Unknown
EcXA181	*ppc*	250	443	4148026	4150677	<	predicted	asp, glu; phosphoenolpyruvate carboxylase (EC 4.1.1.31)
EcXA182	*tktA*	251	444	3077663	3079654	<	predicted	Transketolase (EC 2.2.1.1)
EcXA183	*yabB*	252	445	89634	90092	>	predicted	Unknown
	yabC	253	446	90094	91035	>		Unknown

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
	ftsL	254	447	91032	91397	>		sensitiveCell division and growth; essential gene; cytoplasmic membrane protein
	ftsI	255	448	91413	93179	>		sensitivepbpB, sep; peptidoglycan synthetase; penicillin-binding protein 3
	murE	256	449	93166	94653	>		meso-Diaminopimelate adding enzyme
	murF	257	450	94650	96008	>		mra; D-alanyl:D-alanine adding enzyme
	mraY	258	451	96002	97084	>		UDP-N-acetylmuramoyl-pentapeptide:undecaprenyl-PO4 phosphatase (EC 2.7.8.13)
	murD	259	452	97087	98403	>		UDP-N-acetylmuramoyl-L-alanine:D-glutamate ligase (EC 6.3.2.9)
	ftsW	260	453	98403	99647	>		sensitiveCytoplasmic membrane required for PBP 2 expression; homology to rodA
	murG	261	454	99644	100711	>		UDP-NAc-glucosamine: NAc-muramyl-(pentapeptide) pyrophosphoryl-undecaprenolNAc-glucosamine transferase
	murC	262	455	100765	102240	>		L-Alanine adding enzyme
	ddlB	263	456	102233	103153	>		ddl; D-Alanine:D-alanine ligase
EcXA184	*cysC*	264	457	2871410	2872015	<	documented	Adenylylsulfate kinase (EC 2.7.1.25)
	cysN	265	458	2872015	2873442	<		ATP sulfurylase (ATP:sulfate adenylyltransferase)
EcXA185	*rplQ*	266	459	3437253	3437636	<	documented	50S ribosomal subunit protein L17
	rpoA	267	460	3437677	3438666	<		phs, sez; phage P2 vir1 resistance; RNA polymerase, -subunit (EC 2.7.7.6).
	rpsD	268	461	3438692	3439312	<		ramA, sud2; 30S ribosomal subunit protein S4
	rpsK	269	462	3439346	3439735	<		30S ribosomal subunit protein S11
	rpsM	270	463	3439752	3440108	<		30S ribosomal subunit protein S13
EcXA186	*yagG*	271	464	284619	286001	>	predicted	Unknown

Molecule Number`	Gene	Seq ID (Gene)	Seq ID (Protein)	Start	Stop	Orientation	operon prediction	Predicted function
	b0271	272	465	286013	287623	>		Unknown
EcXA187	*b2271*	273	466	2383874	2384851	>	predicted	Unknown
EcXA188	*b1113*	274	467	1168635	1169597	<	predicted	Unknown
EcXA189a-c	*b1828*	275	468	1908189	1909673	>	predicted	Unknown
EcXA190a-b	*b3052*	276	469	3192961	3194394	<	predicted	Unknown
	glnE	277	470	3194442	3197282	<		GS adenylyl transferase (EC 2.7.7.42)
	ygiF	278	471	3197305	3198606	<		Unknown
EcXA191a-b	*gidB*	279	472	3920685	3921308	<	predicted	Glucose effects on cell division, perhaps replication
	gidA	280	473	3921372	3923261	<		Glucose effects on cell division, perhaps replication
	mioC	281	474	3923640	3924083	<		Initiation of replication; transcription of 16-kDa protein proceeds through ori
EcXA192	*rbsD*	282	475	3930941	3931396	>	documented	rbsP; D-ribose high-affinity transport system
	rbsA	283	476	3931404	3932909	>		rbsP, rbsT; D-ribose high-affinity transport system (may have chemotaxis function)
	rbsC	284	477	3932914	3933879	>		rbsP, rbsT; D-ribose high-affinity transport system
	rbsB	285	478	3933904	3934794	>		prlB, rbsP; D-ribose periplasmic binding protein
	rbsK	286	479	3934920	3935849	>		Ribokinase (EC 2.7.1.15)

Functions for the identified genes were assigned using either Blattner functional class designations, functions referenced in Berlyn, MKB "Linkage Map of *Escherichia coli* K-12, Edition 10: The Traditional Map". (1998) Microbiol. Mol. Biol. Rev. September 62 (3): 814-984, or by comparing identified sequence with known sequences in various databases. A variety of biological functions were noted for the genes to which the clones of the present invention correspond. Biological functions for genes that lie on the same operon as an identified gene have also been made. The functions for the genes of interest appear in Table II.

The genes of interest have a variety of biological functions. For example, genes that are thought to function as transport or binding proteins, that participate in translation or post-translational modification, that are involved in carbon compound catabolism, that are thought to be enzymes, participate in cell processes, energy metabolism and biosynthetic functions appear in Table II. Genes that are involved in cell structure, transcription, RNA processing and degradation also appear in Table II.

Several of the expression vectors contain fragments that correspond to genes of unknown function or if the function is known, it is not known whether the gene is essential. For example, EcXA119, 120, 121, 122a-d, 123, 125, 126, 127a-b, 128, 129, 131, 132, 138, 139a-b, 141, 143, 146, 147, 14, 149a-b, 152, 153, 154, 155, 156, 158, 159, 160, 162, 163, 164, 165, 166, 167, 168, 170, 171, 172, 173, 176, 177, 180, 181, 186, 187, 188, 189a-b, 190a-b, 191a-b, and 192 are all exogenous nucleic acid sequences that correspond to *E. coli* proteins that have no known function or where the function has not been shown to be essential or nonessential.

The present invention reports a number of novel *E. coli* genes and operons that are required for proliferation. From the list of clone sequences identified here, each was identified to be a portion of a gene in an operon required for the proliferation of *E. coli*. Cloned sequences corresponding to genes already known to be required for proliferation in *E. coli* include EcXA118a-d, 124, 130, 133a-c, 136a-b, 142, 145, 150, 157, 169, 178, 182, 183 and 185 are exogenous nucleic acid sequences that correspond to *E. coli* genes that are known to be required for cellular proliferation. The remaining identified sequences correspond to *E. coli* genes previously undesignated as required for proliferation in the art.

An interesting observation of the present invention is that there are also several sequence fragments that correspond to *E. coli* genes that are not thought to be required for *E. coli* proliferation. Nevertheless, under the conditions described above, the antisense expression of these gene fragments causes a reduction in cell growth. This result implies that the genes corresponding to the identified sequences are actually required for proliferation or are in operons required for proliferation. Molecule Nos. corresponding to these genes are EcXA128, 134, 137, 140a-b, 144a-c, 151, 161, 174, 175, and 184.

Following identification of the sequences of interest, these sequences were localized into operons. Since bacterial genes are expressed in a polycistronic manner, the antisense inhibition of a

single gene in an operon might effect the expression of all the other genes on the operon or the genes down stream from the single gene identified. In order to determine which of the gene products in an operon are required for proliferation, each of the genes contained within an operon may be analyzed for their effect on viability as described below.

TABLE III

Operon Boundaries

Molecule Number	Start	Stop
EcXA118a-d	3440255	3445786
EcXA119	3022315	3026508
EcXA120	1590689	1596110
EcXA121	4415276	4416031
EcXA122	2183937	2190242
EcXA123	3472315	3474089
EcXA124	3445951	3450907
EcXA125	4090705	4091019
EcXA126	3814303	3815166
EcXA127a-b	2989290	2989781
EcXA128	703167	705113
EcXA129	149715	152854
EcXA130	3913181	3920080
EcXA131	3156944	3159162
EcXA132	1533961	1540614
EcXA133a-c	Same operon as EcXA118	
EcXA134	3373914	3375055
EcXA136a-b	3031677	3037763
EcXA137	259612	261980
EcXA138	3265474	3267243
EcXA139a-b	913181	914128
EcXA140a-b	3848429	3850615
EcXA141	2099917	2106351
EcXA142	Same operon as EcXA118	
EcXA143	2985498	2986190
EcXA144a-c	1686600	1687775
EcXA145	2962383	2964059
EcXA146	3904481	3909153
EcXA147	1167423	1168133
EcXA148	3972208	3976217
EcXA149a-b	3054261	3054809
EcXA150	3819733	3824918
EcXA151	3679791	3681077
EcXA152	2276590	2278503
EcXA153	4248534	4249862
EcXA154	2054880	2055596
EcXA155	2652177	2652962
EcXA156	4059826	4061091
EcXA157	Same operon as	

Molecule Number	Start	Stop
	EcXA124	
EcXA158	2643033	2645345
	2645346	2650307
EcXA159	506510	507304
EcXA160	1823979	1830006
EcXA161	411831	416176
EcXA162	4490155	4491974
EcXA163	3882705	3884351
EcXA164	2350667	2354731
EcXA165	686062	686970
EcXA166	4039996	4041625
EcXA167	2169417	2175230
EcXA168	1801118	1803017
EcXA169	Same operon as EcXA118	
EcXA170	381728	383159
EcXA171	915696	917354
EcXA172	2461032	2462090
EcXA173	3623310	3628232
EcXA174	3088366	3089073
EcXA175	1935673	1937115
EcXA176	4346893	4349240
EcXA177	2493599	2494585
EcXA178	20815	21078
EcXA180	4279362	4280654
EcXA181	4148026	4150677
EcXA182	3077663	3079654
EcXA183	89634	103153
EcXA184	2871410	2873442
EcXA185	3437253	3440108
EcXA186	284619	287623
EcXA187	2383874	2384851
EcXA188	1168635	1169597
EcXA189a-c	1908189	1909673
EcXA190a-b	3192961	3198606
EcXA191a-b	3920685	3924083
EcXA192	3930941	3935849

EXAMPLE 5

Identification of Individual Genes within an Operon Required for Proliferation

The following example illustrates a method for determining which gene in an operon is required for proliferation. The clone insert corresponding to Molecule No. EcXA119 possesses nucleic acid sequence homology to the *E. coli* gene *b2883*. This gene is located in an operon containing the *b2882, b2883, b2884,* and *b2885* genes. To determine which gene or genes in this operon are required for proliferation, each gene is selectively inactivated using homologous recombination. Gene *b2885* is the first gene to be inactivated.

Deletion inactivation of a chromosomal copy of a gene in *E. coli* can be accomplished by integrative gene replacement. The principle of this method (Hamilton, C. M., et al 1989. *J. Bacteriol.* **171**: 4617-4622) is to construct a mutant allele of the targeted gene, introduce that allele into the chromosome using a conditional suicide vector, and then force the removal of the native wild type allele and vector sequences. This will replace the native gene with a desired mutation(s) but leave promoters, operators, etc. intact. Essentiality of a gene is determined either by deduction from genetic analysis or by conditional expression of a wild type copy of the targeted gene (trans complementation).

The first step is to generate a mutant *b2885* allele using PCR amplification. Two sets of PCR primers are chosen to produce a copy of *b2885* with a large central deletion to inactivate the gene. In order to eliminate polar effects, it is desirable to construct a mutant allele comprising an in-frame deletion of most or all of the coding region of the *b2885* gene. Each set of PCR primers is chosen such that a region flanking the gene to be amplified is sufficiently long to allow recombination (typically at least 500 nucleotides on each side of the deletion). The targeted deletion or mutation will be contained within this fragment. To facilitate cloning of the PCR product, the PCR primers may also contain restriction endonuclease sites found in the cloning region of a conditional knockout vector such as pKO3 (Link, et al 1997 *J. Bacteriol.* **179** (20): 6228-6237). Suitable sites include *NotI*, *SalI*, *Bam*HI and *SmaI*. The b2885 gene fragments are produced using standard PCR conditions including, but not limited to, those outlined in the manufacturers directions for the Hot Start Taq PCR kit (Qiagen, Inc., Valencia, CA). The PCR reactions will produce two fragments that can be fused together. Alternatively, crossover PCR can be used to generate a desired deletion in one step (Ho, S. N., et al 1989. *Gene* **77**: 51-59, Horton, R. M., et al 1989. *Gene* **77**: 61-68). The mutant allele thus produced is called a "null" allele because it cannot produce a functional gene product.

The mutant allele obtained from PCR amplification is cloned into the multiple cloning site of pKO3. Directional cloning of the *b2885* null allele is not necessary. The pKO3 vector has a temperature-sensitive origin of replication derived from pSC101. Therefore, clones are propagated at the permissive temperature of 30°C. The vector also contains two selectable marker genes: one that confers resistance to chloramphenicol and another, the *Bacillus subtilis sacB* gene, that allows for counter-selection on sucrose containing growth medium. Clones that contain vector DNA with the null allele inserted are confirmed by restriction endonuclease analysis and DNA sequence analysis of isolated plasmid DNA. The plasmid containing the *b2885* null allele insert is known as a knockout plasmid.

Once the knockout plasmid has been constructed and its sequence verified, it is transformed into a Rec^+ *E. coli* host cell. Transformation can be by any standard method such as electroporation. In some fraction of the transformed cells, plasmids will integrate into the *E. coli* chromosome by homologous recombination between the *b2885* null allele in the plasmid and the

b2885 gene in the chromosome. Transformant colonies in which such an event has occurred are readily selected by growth at the non-permissive temperature of 43°C and in the presence of choramphenicol. At this temperature, the plasmid will not replicate as an episome and will be lost from cells as they grow and divide. These cells are no longer resistant to chloramphenicol and will not grow when it is present. However, cells in which the knockout plasmid has integrated into the *E. coli* chromosome remain resistant to chloramphenicol and propagate.

Cells containing integrated knock-out plasmids are usually the result of a single crossover event that creates a tandem repeat of the mutant and native wild type alleles of *b2885* separated by the vector sequences. A consequence of this is that *b2885* will still be expressed in these cells. In order to determine if the gene is essential for growth, the wild type copy must be removed. This is accomplished by selecting for plasmid excision, a process in which homologous recombination between the two alleles results in looping out of the plasmid sequences. Cells that have undergone such an excision event and have lost plasmid sequences including *sacB* gene are selected for by addition of sucrose to the medium. The *sacB* gene product converts sucrose to a toxic molecule. Thus counter selection with sucrose ensures that plasmid sequences are no longer present in the cell. Loss of plasmid sequences is further confirmed by testing for sensitivity to chloramphenicol (loss of the chloramphenicol resistance gene). The latter test is important because occasionally a mutation in the *sacB* gene can occur resulting in a loss of *sacB* function with no effect on plasmid replication (Link, et. al., 1997 *J. Bacteriol.* **179** (20): 6228-6237). These artifact clones retain plasmid sequences and are therefore still resistant to chloramphenicol.

In the process of plasmid excision, one of the two *b2885* alleles is lost from the chromosome along with the plasmid DNA. In general, it is equally likely that the null allele or the wild type allele will be lost. Therefore, if the *b2885* gene is not essential, half of the clones obtained in this experiment will have the wild type allele on the chromosome and half will have the null allele. However, if the *b2885* gene is essential, cells containing the null allele will not be obtained as a single copy of the null allele would be lethal.

To determine the essentiality of *b2885,* a statistically significant number of the resulting clones, at least 20, are analyzed by PCR amplification of the b2885 gene. Since the null allele is missing a significant portion of the *b2885* gene, its PCR product is significantly shorter than that of the wild type gene and the two are readily distinguished by gel electrophoretic analysis. The PCR products may also be subjected to sequence determination for further confirmation by methods well known to those in the art.

The above experiment is generally adequate for determining the essentiality of a gene such as *b2885*. However, it may be necessary or desirable to more directly confirm the essentiality of the gene. There are several methods by which this can be accomplished. In general, these involve three steps: 1) construction of an episome containing a wild type allele, 2) isolation of clones containing a single chromosomal copy of the mutant null allele as described above but in the

presence of the episomal wild type allele, and then 3) determining if the cells survive when the expression of the episomal allele is shut off. In this case, the trans copy of wild type *b2885* is made by PCR cloning of the entire coding region of *b2885* and inserting it in the sense orientation downstream of an inducible promoter such as the *E. coli lac* promoter. Transcription of this allele of *b2885* will be induced in the presence of IPTG which inactivates the *lac* repressor. Under IPTG induction *b2885* protein will be expressed as long as the recombinant gene also possesses a ribosomal binding site, also known as a "Shine-Dalgarno Sequence". The trans copy of *b2885* is cloned on a plasmid that is compatible with pSC101. Compatible vectors include p15A, pBR322, and the pUC plasmids, among others. Replication of the compatible plasmid will not be temperature-sensitive. The entire process of integrating the null allele of *b2885* and subsequent plasmid excision is carried out in the presence of IPTG to ensure the expression of functional *b2885* protein is maintained throughout. After the null *b2885* allele is confirmed as integrated on the chromosome in place of the wild type *b2885* allele, then IPTG is withdrawn and expression of functional *b2885* protein shut off. If the *b2885* gene is essential, cells will cease to proliferate under these conditions. However, if the *b2885* gene is not essential, cells will continue to proliferate under these conditions. In this experiment, essentiality is determined by conditional expression of a wild type copy of the gene rather than inability to obtain the intended chromosomal disruption.

An advantage of this method over some other gene disruption techniques is that the targeted gene can be deleted or mutated without the introduction of large segments of foreign DNA. Therefore, polar effects on downstream genes are eliminated or minimized. There are methods described to introduce inducible promoters upstream of potential essential bacterial genes. However in such cases, polarity from multiple transcription start points can be a problem. One way of preventing this is to insert a gene disruption cassette that contains strong transcriptional terminators upstream of the integrated inducible promoter (Zhang, Y, and Cronan, J. E. 1996 *J. Bacteriol.* **178** (12): 3614-3620). The described techniques will all be familiar to one of ordinary skill in the art.

Following the analysis of the *b2885* gene, the other genes of the operon are investigated to determine if they are required for proliferation.

EXAMPLE 6

Expression of the Proteins Encoded by Genes Identified as Required for *E. coli* Proliferation

The following is provided as one exemplary method to express the proliferation-required proteins encoded by the identified sequences described above. First, the initiation and termination codons for the gene are identified. If desired, methods for improving translation or expression of the protein are well known in the art. For example, if the nucleic acid encoding the polypeptide to be expressed lacks a methionine codon to serve as the initiation site, a strong Shine-Delgarno sequence, or a stop codon, these sequences can be added. Similarly, if the identified nucleic acid sequence lacks a

transcription termination signal, this sequence can be added to the construct by, for example, splicing out such a sequence from an appropriate donor sequence. In addition, the coding sequence may be operably linked to a strong promoter or an inducible promoter if desired. The identified nucleic acid sequence or portion thereof encoding the polypeptide to be expressed is obtained by PCR from the bacterial expression vector or genome using oligonucleotide primers complementary to the identified nucleic acid sequence or portion thereof and containing restriction endonuclease sequences for *Nco*I incorporated into the 5' primer and *Bgl*II at the 5' end of the corresponding 3'-primer, taking care to ensure that the identified nucleic acid sequence is positioned in frame with the termination signal. The purified fragment obtained from the resulting PCR reaction is digested with *Nco*I and *Bgl*II, purified and ligated to an expression vector.

The ligated product is transformed into DH5α or some other *E. coli* strain suitable for the over expression of potential proteins. Transformation protocols are well known in the art. For example, transformation protocols are described in: **Current Protocols in Molecular Biology,** Vol. 1, Unit 1.8, (Ausubel, et al., Eds.) John Wiley & Sons, Inc. (1997). Positive transformants are selected after growing the transformed cells on plates containing 50-100 μg/ml Ampicillin (Sigma, St. Louis, Missouri). In one embodiment, the expressed protein is held in the cytoplasm of the host organism. In an alternate embodiment, the expressed protein is released into the culture medium. In still another alternative, the expressed protein can be sequestered in the periplasmic space and liberated therefrom using any one of a number of cell lysis techniques known in the art. For example, the osmotic shock cell lysis method described in Chapter 16 of **Current Protocols in Molecular Biology,** Vol. 2, (Ausubel, et al., Eds.) John Wiley & Sons, Inc. (1997). Each of these procedures can be used to express a proliferation-required protein.

Expressed proteins, whether in the culture medium or liberated from the periplasmic space or the cytoplasm, are then purified or enriched from the supernatant using conventional techniques such as ammonium sulfate precipitation, standard chromatography, immunoprecipitation, immunochromatography, size exclusion chromatography, ion exchange chromatography, and HPLC. Alternatively, the secreted protein can be in a sufficiently enriched or pure state in the supernatant or growth media of the host to permit it to be used for its intended purpose without further enrichment. The purity of the protein product obtained can be assessed using techniques such as Coomassie or silver staining or using antibodies against the control protein. Coomassie and silver staining techniques are familiar to those skilled in the art.

Antibodies capable of specifically recognizing the protein of interest can be generated using synthetic peptides using methods well known in the art. See, **Antibodies: A Laboratory Manual,** (Harlow and Lane, Eds.) Cold Spring Harbor Laboratory (1988). For example, 15-mer peptides having a sequence encoded by the appropriate identified gene sequence of interest or portion thereof can be chemically synthesized. The synthetic peptides are injected into mice to generate antibodies to the polypeptide encoded by the identified nucleic acid sequence of interest or portion thereof.

Alternatively, samples of the protein expressed from the expression vectors discussed above can be purified and subjected to amino acid sequencing analysis to confirm the identity of the recombinantly expressed protein and subsequently used to raise antibodies. An Example describing in detail the generation of monoclonal and polyclonal antibodies appears in Example 7.

The protein encoded by the identified nucleic acid sequence of interest or portion thereof can be purified using standard immunochromatography techniques. In such procedures, a solution containing the secreted protein, such as the culture medium or a cell extract, is applied to a column having antibodies against the secreted protein attached to the chromatography matrix. The secreted protein is allowed to bind the immunochromatography column. Thereafter, the column is washed to remove non-specifically bound proteins. The specifically bound secreted protein is then released from the column and recovered using standard techniques. These procedures are well known in the art.

In an alternative protein purification scheme, the identified nucleic acid sequence of interest or portion thereof can be incorporated into expression vectors designed for use in purification schemes employing chimeric polypeptides. In such strategies the coding sequence of the identified nucleic acid sequence of interest or portion thereof is inserted in-frame with the gene encoding the other half of the chimera. The other half of the chimera can be maltose binding protein (MBP) or a nickel binding polypeptide encoding sequence. A chromatography matrix having antibody to MBP or nickel attached thereto is then used to purify the chimeric protein. Protease cleavage sites can be engineered between the MBP gene or the nickel binding polypeptide and the identified expected gene of interest, or portion thereof. Thus, the two polypeptides of the chimera can be separated from one another by protease digestion.

One useful expression vector for generating maltose binding protein fusion proteins is pMAL (New England Biolabs), which encodes the *malE* gene. In the pMal protein fusion system, the cloned gene is inserted into a pMal vector downstream from the *malE* gene. This results in the expression of an MBP-fusion protein. The fusion protein is purified by affinity chromatography. These techniques as described are well known to those skilled in the art of molecular biology.

EXAMPLE 7

Production of an Antibody to an isolated *E. coli* Protein

Substantially pure protein or polypeptide is isolated from the transformed cells as described in Example 6. The concentration of protein in the final preparation is adjusted, for example, by concentration on a 10,000 molecular weight cut off AMICON filter device (Millipore, Bedford, MA), to the level of a few micrograms/ml. Monoclonal or polyclonal antibody to the protein can then be prepared as follows:

Monoclonal Antibody Production by Hybridoma Fusion

Monoclonal antibody to epitopes of any of the peptides identified and isolated as described can be prepared from murine hybridomas according to the classical method of Kohler, G. and Milstein, C., **Nature 256**:495 (1975) or any of the well-known derivative methods thereof. Briefly, a mouse is repetitively inoculated with a few micrograms of the selected protein or peptides derived therefrom

over a period of a few weeks. The mouse is then sacrificed, and the antibody producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and the excess unfused cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). The successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate where growth of the culture is continued. Antibody-producing clones are identified by detection of antibody in the supernatant fluid of the wells by immunoassay procedures, such as ELISA, as described by Engvall, E., "Enzyme immunoassay ELISA and EMIT," **Meth. Enzymol. 70**:419 (1980), and derivative methods thereof. Selected positive clones can be expanded and their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Davis, L. et al. **Basic Methods in Molecular Biology** Elsevier, New York. Section 21-2.

Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogeneous epitopes of a single protein or a peptide can be prepared by immunizing suitable animals with the expressed protein or peptides derived therefrom described above, which can be unmodified or modified to enhance immunogenicity. Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. For example, small molecules tend to be less immunogenic than larger molecules and can require the use of carriers and adjuvant. Also, host animals vary in response to site of inoculations and dose, with both inadequate or excessive doses of antigen resulting in low titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appears to be most reliable. An effective immunization protocol for rabbits can be found in Vaitukaitis, J. et al. **J. Clin. Endocrinol. Metab. 33**:988-991 (1971).

Booster injections can be given at regular intervals, and antiserum harvested when antibody titer thereof, as determined semi-quantitatively, for example, by double immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony, O. et al., Chap. 19 in: **Handbook of Experimental Immunology** D. Wier (ed) Blackwell (1973). Plateau concentration of antibody is usually in the range of 0.1 to 0.2 mg/ml of serum (about 12μM). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher, D., Chap. 42 in: **Manual of Clinical Immunology**, 2d Ed. (Rose and Friedman, Eds.) Amer. Soc. For Microbiol., Washington, D.C. (1980).

Antibody preparations prepared according to either protocol are useful in quantitative immunoassays which determine concentrations of antigen-bearing substances in biological samples; they are also used semi-quantitatively or qualitatively to identify the presence of antigen in a biological sample. The antibodies can also be used in therapeutic compositions for killing bacterial cells expressing the protein.

EXAMPLE 8

Screening Chemical Libraries

A. Protein-Based Assays

Having isolated and expressed bacterial proteins shown to be required for bacterial proliferation, the present invention further contemplates the use of these expressed proteins in assays to screen libraries of compounds for potential drug candidates. The generation of chemical libraries is well known in the art. For example combinatorial chemistry can be used to generate a library of compounds to be screened in the assays described herein. A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical "building blocks" reagents. For example, a linear combinatorial chemical library such as a polypeptide library is formed by combining amino acids in every possible combination to yield peptides of a given length. Millions of chemical compounds theoretically can be synthesized through such combinatorial use of chemical building blocks. For example, one commentator observed that the systematic, combinatorial mixing of 100 interchangeable chemical building blocks results in the theoretical synthesis of 100 million tetrameric compounds or 10 billion pentameric compounds. (Gallop et al., "Applications of Combinatorial Technologies to Drug Discovery, Background and Peptide Combinatorial Libraries," **Journal of Medicinal Chemistry**, Vol. 37, No. 9, 1233-1250 (1994). Other chemical libraries known to those in the art may also be used, including natural product libraries.

Once generated, combinatorial libraries can be screened for compounds that possess desirable biological properties. For example, compounds which may be useful as drugs or to develop drugs would likely have the ability to bind to the target protein identified, expressed and purified as discussed above. Further, if the identified target protein is an enzyme, candidate compounds would likely interfere with the enzymatic properties of the target protein. Any enzyme can be a target protein. For example, the enzymatic function of a target protein can be to serve as a protease, nuclease, phosphatase, dehydrogenase, transporter protein, transcriptional enzyme, and any other type of enzyme known or unknown. Thus, the present invention contemplates using the protein products described above to screen combinatorial and other chemical libraries.

Those in the art will appreciate that a number of techniques exist for characterizing target proteins in order to identify molecules useful for the discovery and development of therapeutics. For example, some techniques involve the generation and use of small peptides to probe and analyze target proteins both biochemically and genetically in order to identify and develop drug leads. Such techniques include the methods described in PCT publications No. WO9935494, WO9819162, WO9954728.

In another example, the target protein is a serine protease and the substrate of the enzyme is known. The present example is directed towards the analysis of libraries of compounds to identify compounds that function as inhibitors of the target enzyme. First, a library of small molecules is

generated using methods of combinatorial library formation well known in the art. U.S. Patent NOs. 5,463,564 and 5,574, 656, to Agrafiotis, et al., entitled "System and Method of Automatically Generating Chemical Compound with Desired Properties," are two such teachings. Then the library compounds are screened to identify library compounds that possess desired structural and functional properties. U.S. Patent No. 5,684,711, also discusses a method for screening libraries.

To illustrate the screening process, the combined target and chemical compounds of the library are exposed to and permitted to interact with the purified enzyme. A labeled substrate is added to the incubation. The label on the substrate is such that a detectable signal is emitted from metabolized substrate molecules. The emission of this signal permits one to measure the effect of the combinatorial library compounds on the enzymatic activity of target enzymes. The characteristics of each library compound is encoded so that compounds demonstrating activity against the enzyme can be analyzed and features common to the various compounds identified can be isolated and combined into future iterations of libraries.

Once a library of compounds is screened, subsequent libraries are generated using those chemical building blocks that possess the features shown in the first round of screen to have activity against the target enzyme. Using this method, subsequent iterations of candidate compounds will possess more and more of those structural and functional features required to inhibit the function of the target enzyme, until a group of enzyme inhibitors with high specificity for the enzyme can be found. These compounds can then be further tested for their safety and efficacy as antibiotics for use in mammals.

It will be readily appreciated that this particular screening methodology is exemplary only. Other methods are well known to those skilled in the art. For example, a wide variety of screening techniques are known for a large number of naturally-occurring targets when the biochemical function of the target protein is known.

B. Cell-based Assays

Current cell-based assays used to identify or to characterize compounds for drug discovery and development frequently depend on detecting the ability of a test compound to inhibit the activity of a target molecule located within a cell or located on the surface of a cell. An advantage of cell-based assays is that they allow the effect of a compound on a target molecule's activity to be detected within the physiologically relevant environment of the cell as opposed to an *in vitro* environment. Most often such target molecules are proteins such as enzymes, receptors and the like. However, target molecules may also include other molecules such as DNAs, lipids, carbohydrates and RNAs including messenger RNAs, ribosomal RNAs, tRNAs and the like. A number of highly sensitive cell-based assay methods are available to those of skill in the art to detect binding and interaction of test compounds with specific target molecules. However, these methods are generally not highly effective when the test compound binds to or otherwise interacts with its target molecule with moderate or low affinity. In addition, the target molecule may not be

readily accessible to a test compound in solution, such as when the target molecule is located inside the cell or within a cellular compartment such as the periplasm of a bacterial cell. Thus, current cell-based assay methods are limited in that they are not effective in identifying or characterizing compounds that interact with their targets with moderate to low affinity or compounds that interact with targets that are not readily accessible.

Cell-based assay methods of the present invention have substantial advantages over current cell-based assays practiced in the art. These advantages derive from the use of sensitized cells in which the level or activity of a proliferation-required gene product (the target molecule) has been specifically reduced to the point where the presence or absence of its function becomes a rate-determining step for cellular proliferation. Bacterial, fungal, plant, or animal cells can all be used with the present method. Such sensitized cells become much more sensitive to compounds that are active against the affected target molecule. Thus, cell-based assays of the present invention are capable of detecting compounds exhibiting low or moderate potency against the target molecule of interest because such compounds are substantially more potent on sensitized cells than on non-sensitized cells. The affect may be such that a test compound may be two to several times more potent, at least 10 times more potent, at least 20 times more potent, at least 50 times more potent, at least 100 times more potent, at least 1000 times more potent, or even more than 1000 times more potent when tested on the sensitized cells as compared to the non-sensitized cells.

Due in part to the increased appearance of antibiotic resistance in pathogenic microorganisms and to the significant side-effects associated with some currently used antibiotics, novel antibiotics acting at new targets are highly sought after in the art. Yet, another limitation in the current art related to cell-based assays is the problem of identifying hits against the same kinds of target molecules in the same limited set of biological pathways over and over again. This may occur when compounds acting at such new targets are discarded, ignored or fail to be detected because compounds acting at the "old" targets are encountered more frequently and are more potent than compounds acting at the new targets. As a result, the majority of antibiotics in use currently interact with a relatively small number of target molecules within an even more limited set of biological pathways.

The use of sensitized cells of the current invention provides a solution to the above problem in two ways. First, desired compounds acting at a target of interest, whether a new target or a previously known but poorly exploited target, can now be detected above the "noise" of compounds acting at the "old" targets due to the specific and substantial increase in potency of such desired compounds when tested on the sensitized cells of the current invention. Second, the methods used to sensitize cells to compounds acting at a target of interest may also sensitize these cells to compounds acting at other target molecules within the same biological pathway. For example, expression of an antisense molecule to a gene encoding a ribosomal protein is expected to sensitize the cell to compounds acting at that ribosomal protein and may also sensitize the cells to

compounds acting at any of the ribosomal components (proteins or rRNA) or even to compounds acting at any target which is part of the protein synthesis pathway. Thus an important advantage of the present invention is the ability to reveal new targets and pathways that were previously not readily accessible to drug discovery methods.

Sensitized cells of the present invention are prepared by reducing the activity or level of a target molecule. The target molecule may be a gene product, such as an RNA or polypeptide produced from the proliferation-required nucleic acids described herein. Alternatively, the target may be a gene product such as an RNA or polypeptide which is produced from a sequence within the same operon as the proliferation-required nucleic acids described herein. In addition, the target may be an RNA or polypeptide in the same biological pathway as the proliferation-required nucleic acids described herein. Such biological pathways include, but are not limited to, enzymatic, biochemical and metabolic pathways as well as pathways involved in the production of cellular structures such the cell wall.

Current methods employed in the arts of medicinal and combinatorial chemistries are able to make use of structure-activity relationship information derived from testing compounds in various biological assays including direct binding assays and cell-based assays. Occasionally compounds are directly identified in such assays that are sufficiently potent to be developed as drugs. More often, initial hit compounds exhibit moderate or low potency. Once a hit compound is identified with low or moderate potency, directed libraries of compounds are synthesized and tested in order to identify more potent leads. Generally these directed libraries are combinatorial chemical libraries consisting of compounds with structures related to the hit compound but containing systematic variations including additions, subtractions and substitutions of various structural features. When tested for activity against the target molecule, structural features are identified that either alone or in combination with other features enhance or reduce activity. This information is used to design subsequent directed libraries containing compounds with enhanced activity against the target molecule. After one or several iterations of this process, compounds with substantially increased activity against the target molecule are identified and may be further developed as drugs. This process is facilitated by use of the sensitized cells of the present invention since compounds acting at the selected targets exhibit increased potency in such cell-based assays, thus; more compounds can now be characterized providing more useful information than would be obtained otherwise.

Thus, it is now possible using cell-based assays of the present invention to identify or characterize compounds that previously would not have been readily identified or characterized including compounds that act at targets that previously were not readily exploited using cell-based assays. The process of evolving potent drug leads from initial hit compounds is also substantially improved by the cell-based assays of the present invention because, for the same number of test compounds, more structure-function relationship information is likely to be revealed.

The method of sensitizing a cell entails selecting a suitable gene or operon. A suitable gene or operon is one whose expression is required for the proliferation of the cell to be sensitized. The next step is to introduce into the cells to be sensitized, an antisense RNA capable of hybridizing to the suitable gene or operon or to the RNA encoded by the suitable gene or operon. Introduction of the antisense RNA can be in the form of an expression vector in which antisense RNA is produced under the control of an inducible promoter. The amount of antisense RNA produced is limited by varying the inducer concentration to which the cell is exposed and thereby varying the activity of the promoter driving transcription of the antisense RNA. Thus, cells are sensitized by exposing them to an inducer concentration that results in a sub-lethal level of antisense RNA expression.

In one embodiment of the cell-based assays, the identified exogenous *E. coli* nucleotide sequences of the present invention are used to inhibit the production of a proliferation-required protein. Expression vectors producing antisense RNA complementary to identified genes required for proliferation are used to limit the concentration of a proliferation-required protein without severely inhibiting growth. To achieve that goal, a growth inhibition dose curve of inducer is calculated by plotting various doses of inducer against the corresponding growth inhibition caused by the antisense expression. From this curve, various percentages of antisense induced growth inhibition, from 1 to 100% can be determined. If the promoter contained in the expression vector contains a *lac* operator the transcription is regulated by *lac* repressor and expression from the promoer is inducible with IPTG. For example, the highest concentration of the inducer IPTG that does not reduce the growth rate significantly (0% growth inhibition) can be predicted from the curve. Cellular proliferation can be monitored by growth medium turbidity via OD measurements. In another example, the concentration of inducer that reduces growth by 25% can be predicted from the curve. In still another example, a concentration of inducer that reduces growth by 50% can be calculated. Additional parameters such as colony forming units (cfu) can be used to measure cellular viability.

Cells to be assayed are exposed to the above-determined concentrations of inducer. The presence of the inducer at this sub-lethal concentration reduces the amount of the proliferation required gene product to a low amount in the cell that will limit but not prevent growth. Cells grown in the presence of this concentration of inducer are therefore specifically more sensitive to inhibitors of the proliferation-required protein or RNA of interest or to inhibitors of proteins or RNAs in the same biological pathway as the proliferation-required protein or RNA of interest but not to inhibitors of unrelated proteins or RNAs.

Cells pretreated with sub-inhibitory concentrations of inducer and thus containing a reduced amount of proliferation-required target gene product are then used to screen for compounds that reduce cell growth. The sub-lethal concentration of inducer may be any concentration consistent with the intended use of the assay to identify candidate compounds to which the cells are more sensitive. For example, the sub-lethal concentration of the inducer may be such that growth

inhibition is at least about 5%, at least about 8%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60% at least about 75%, 90%, 95% or more. Cells which are pre-sensitized using the preceding method are more sensitive to inhibitors of the target protein because these cells contain less target protein to be inhibited than do wild-type cells.

In another embodiment of the cell-based assays of the present invention, the level or activity of a proliferation required gene product is reduced using a mutation, such as a temperature sensitive mutation, in the proliferation-required sequence and an antisense nucleic acid complementary to the proliferation-required sequence. Growing the cells at an intermediate temperature between the permissive and restrictive temperatures of the temperature sensitive mutant where the mutation is in a proliferation-required gene produces cells with reduced activity of the proliferation-required gene product. The antisense RNA complementary to the proliferation-required sequence further reduces the activity of the proliferation required gene product. Drugs that may not have been found using either the temperature sensitive mutation or the antisense nucleic acid alone may be identified by determining whether cells in which expression of the antisense nucleic acid has been induced and which are grown at a temperature between the permissive temperature and the restrictive temperature are substantially more sensitive to a test compound than cells in which expression of the antisense nucleic acid has not been induced and which are grown at a permissive temperature. Also drugs found previously from either the antisense nucleic acid alone or the temperature sensitive mutation alone may have a different sensitivity profile when used in cells combining the two approaches, and that sensitivity profile may indicate a more specific action of the drug in inhibiting one or more activities of the gene product.

Temperature sensitive mutations may be located at different sites within the gene and correspond to different domains of the protein. For example, the *dnaB* gene of *Escherichia coli* encodes the replication fork DNA helicase. DnaB has several domains, including domains for oligomerization, ATP hydrolysis, DNA binding, interaction with primase, interaction with DnaC, and interaction with DnaA [(Biswas, E.E. and Biswas, S.B. 1999 Mechanism and DnaB helicase of *Escherichia coli*: structural domains involved in ATP hydrolysis, DNA binding, and oligomerization. Biochem. 38:10919-10928; Hiasa, H. and Marians, K.J. 1999 Initiation of bidirectional replication at the chromosomal origin is directed by the interaction between helicase and primase. J. Biol. Chem. 274:27244-27248; San Martin, C., Radermacher, M., Wolpensinger, B., Engel, A., Miles, C.S., Dixon, N.E., and Carazo, J.M. 1998 Three-dimensional reconstructions from cryoelectron microscopy images reveal an intimate complex between helicase DnaB and its loading partner DnaC. Structure 6:501-9; Sutton, M.D., Carr, K.M., Vicente, M., and Kaguni, J.M. 1998 *Escherichia coli* DnaA protein. The N-terminal domain and loading of DnaB helicase at the *E. coli* chromosomal. J. Biol. Chem. 273:34255-62)]. Temperature sensitive mutations in different domains of DnaB confer different phenotypes at the restrictive temperature, which include either an

abrupt stop or slow stop in DNA replication with or without DNA breakdown (Wechsler, J.A. and Gross, J.D. 1971. *Escherichia coli* mutants temperature-sensitive for DNA synthesis. Mol. Gen. Genetics 113:273-284) and termination of growth or cell death. Combining the use of temperature sensitive mutations in the *dnaB* gene that cause cell death at the restrictive temperature with an antisense to the *dnaB* gene could lead to the discovery of very specific and effective inhibitors of one or a subset of activities exhibited by DnaB.

It will be appreciated that the above method may be performed with any mutation which reduces but does not eliminate the activity or level of the gene product which is required for proliferation.

When screening for antimicrobial agents against a gene product required for proliferation, growth inhibition of cells containing a limiting amount of that proliferation-required gene product can be assayed. Growth inhibition can be measured by directly comparing the amount of growth, measured by the optical density of the growth medium, between an experimental sample and a control sample. Alternative methods for assaying cell proliferation include measuring green fluorescent protein (GFP) reporter construct emissions, various enzymatic activity assays, and other methods well known in the art.

It will be appreciated that the above method may be performed in solid phase, liquid phase or a combination of the two. For example, cells grown on nutrient agar containing the inducer of the antisense construct may be exposed to compounds spotted onto the agar surface. A compound's effect may be judged from the diameter of the resulting killing zone, the area around the compound application point in which cells do not grow. Multiple compounds may be transferred to agar plates and simultaneously tested using automated and semi-automated equipment including but not restricted to multi-channel pipettes (for example the Beckman Multimek) and multi-channel spotters (for example the Genomic Solutions Flexys). In this way multiple plates and thousands to millions of compounds may be tested per day.

The compounds may also be tested entirely in liquid phase using microtiter plates as described below. Liquid phase screening may be performed in microtiter plates containing 96, 384, 1536 or more wells per microtiter plate to screen multiple plates and thousands to millions of compounds per day. Automated and semi-automated equipment may be used for addition of reagents (for example cells and compounds) and determination of cell density.

EXAMPLE 9

Cell-based Assay Using Antisense Complementary to Genes Encoding Ribosomal Proteins

The effectiveness of the above cell-based assay was validated using constructs expressing antisense RNA to the proliferation required *E. coli* genes *rplL*, *rplJ*, and *rplW* encoding ribosomal proteins L7/L12, L10 and L23 respectively. These proteins are part of the protein synthesis apparatus of the cell and as such are required for proliferation. These constructs were used to test the effect of antisense expression on cell sensitivity to antibiotics known to bind to the ribosome

and thereby inhibit protein synthesis. Constructs expressing antisense RNA to several other genes (*elaD*, *visC*, *yohH*, and *atpE/B*), the products of which are not involved in protein synthesis were used for comparison.

First, pLEX5BA (Krause et al., J. Mol. Biol. 274: 365 (1997)) expression vectors containing antisense constructs to either *rplW* or to *elaD* were introduced into separate *E. coli* cell populations. Vector introduction is a technique well known to those of ordinary skill in the art. The expression vectors of this example contain IPTG inducible promoters that drive the expression of the antisense RNA in the presence of the inducer. However, those skilled in the art will appreciate that other inducible promoters may also be used. Suitable expression vectors are also well known in the art. The *E. coli* antisense clones to genes encoding ribosomal proteins L7/L12, L10 and L23 were used to test the effect of antisense expression on cell sensitivity to the antibiotics known to bind to these proteins. Expression vectors containing antisense to either the genes encoding L7/L12 and L10 or L23 were introduced into separate *E. coli* cell populations.

The cell populations were exposed to a range of IPTG concentrations in liquid medium to obtain the growth inhibitory dose curve for each clone (Fig. 1). First, seed cultures were grown to a particular turbidity that is measured by the optical density (OD) of the growth solution. The OD of the solution is directly related to the number of bacterial cells contained therein. Subsequently, sixteen 200 ul liquid medium cultures were grown in a 96 well microtiter plate at 37° C with a range of IPTG concentrations in duplicate two-fold serial dilutions from 1600 uM to 12.5 uM (final concentration). Additionally, control cells were grown in duplicate without IPTG. These cultures were started from equal amounts of cells derived from the same initial seed culture of a clone of interest. The cells were grown for up to 15 hours and the extent of growth was determined by measuring the optical density of the cultures at 600 nm. When the control culture reached mid-log phase the percent growth (relative to the control culture) for each of the IPTG containing cultures was plotted against the log concentrations of IPTG to produce a growth inhibitory dose response curve for the IPTG. The concentration of IPTG that inhibits cell growth to 50% (IC_{50}) as compared to the 0 mM IPTG control (0% growth inhibition) was then calculated from the curve. Under these conditions, an amount of antisense RNA was produced that reduced the expression levels of *rplW* and *elaD* to a degree such that growth was inhibited by 50%.

Alternative methods of measuring growth are also contemplated. Examples of these methods include measurements of proteins, the expression of which is engineered into the cells being tested and can readily be measured. Examples of such proteins include green fluorescent protein (GFP) and various enzymes.

Cells were pretreated with the selected concentration of IPTG and then used to test the sensitivity of cell populations to tetracycline, erythromycin and other protein synthesis inhibitors. Figure 2 is an IPTG dose response curve in *E. coli* transformed with an IPTG-inducible plasmid containing either an antisense clone to the *E. coli* *rpl*W gene (AS-*rplW*) which encodes ribosomal

protein L23 which is required for protein synthesis and essential for cell proliferation, or an antisense clone to the *elaD* (AS-*elaD*) gene which is not known to be involved in protein synthesis and which is also essential for proliferation.

An example of a tetracycline dose response curve is shown in Figures 2A and 2B for the *rplW* and *elaD* genes, respectively. Cells were grown to log phase and then diluted into media alone or media containing IPTG at concentrations which give 20% and 50% growth inhibition as determined by IPTG dose response curves. After 2.5 hours, the cells were diluted to a final OD_{600} of 0.002 into 96 well plates containing (1) +/- IPTG at the same concentrations used for the 2.5 hour pre-incubation; and (2) serial two-fold dilutions of tetracycline such that the final concentrations of tetracycline range from 1 μg/ml to 15.6 ng/ml and 0 μg/ml. The 96 well plates were incubated at 37°C and the OD_{600} was read by a plate reader every 5 minutes for up to 15 hours. For each IPTG concentration and the no IPTG control, tetracycline dose response curves were determined when the control (absence of tetracycline) reached 0.1 OD_{600}. To compare tetracycline sensitivity with and without IPTG, tetracycline IC_{50s} were determined from the dose response curves (Figs. 3A-B). Cells with reduced levels of L23 (AS-*rplW*) showed increased sensitivity to tetracycline (Fig. 2A) as compared to cells with reduced levels of the *elaD* gene product (AS-*elaD*) (Fig. 2B). Figure 3 shows a summary bar chart in which the ratios of tetracycline IC_{50s} determined in the presence of IPTG which gives 50% growth inhibition versus tetracycline IC_{50s} determined without IPTG (fold increase in tetracycline sensitivity) were plotted. Cells with reduced levels of either L7/L12 (encoded by genes *rplL*, *rplJ*) or L23 (encoded by the *rplW* gene) showed increased sensitivity to tetracycline (Fig. 3). Cells expressing antisense to genes not known to be involved in protein synthesis (AS-*atpB/E*, AS-*visC*, AS-*elaD*, AS-*yohH*) did not show the same increased sensitivity to tetracycline, validating the specificity of this assay (Fig. 3).

In addition to the above, it has been observed in initial experiments that clones expressing antisense RNA to genes involved in protein synthesis (including genes encoding ribosomal proteins L7/L12 & L10, L7/L12 alone, L22, and L18, as well as genes encoding rRNA and Elongation Factor G) have increased sensitivity to the macrolide, erythromycin, whereas clones expressing antisense to the non-protein synthesis genes *elaD*, *atpB/E* and *visC* do not. Furthermore, the clone expressing antisense to *rplL* and *rplJ* does not show increased sensitivity to nalidixic acid and ofloxacin, antibiotics which do not inhibit protein synthesis.

The results with the ribosomal protein genes *rplL*, *rplJ*, and *rplW* as well as the initial results using various other antisense clones and antibiotics show that limiting the concentration of an antibiotic target makes cells more sensitive to the antimicrobial agents that specifically interact with that protein. The results also show that these cells are sensitized to antimicrobial agents that inhibit the overall function in which the protein target is involved but are not sensitized to antimicrobial agents that inhibit other functions.

The cell-based assay described above may also be used to identify the biological pathway in which a proliferation-required nucleic acid or its gene product lies. In such methods, cells expressing a sub-lethal level of antisense to a target proliferation-required nucleic acid and control cells in which expression of the antisense has not been induced are contacted with a panel of antibiotics known to act in various pathways. If the antibiotic acts in the pathway in which the target proliferation-required nucleic acid or its gene product lies, cells in which expression of the antisense has been induced will be more sensitive to the antibiotic than cells in which expression of the antisense has not been induced.

As a control, the results of the assay may be confirmed by contacting a panel of cells expressing antisense nucleic acids to many different proliferation-required genes including the target proliferation-required gene. If the antibiotic is acting specifically, heightened sensitivity to the antibiotic will be observed only in the cells expressing antisense to a target proliferation-required gene (or cells expressing antisense to other proliferation-required genes in the same pathway as the target proliferation-required gene) but will not be observed generally in all cells expressing antisense to proliferation-required genes.

Similarly, the above method may be used to determine the pathway on which a test compound, such as a test antibiotic acts. A panel of cells, each of which expresses antisense to a proliferation-required nucleic acid in a known pathway, is contacted with a compound for which it is desired to determine the pathway on which it acts. The sensitivity of the panel of cells to the test compound is determined in cells in which expression of the antisense has been induced and in control cells in which expression of the antisense has not been induced. If the test compound acts on the pathway on which an antisense nucleic acid acts, cells in which expression of the antisense has been induced will be more sensitive to the compound than cells in which expression of the antisense has not been induced. In addition, control cells in which expression of antisense to proliferation-required genes in other pathways has been induced will not exhibit heightened sensitivity to the compound. In this way, the pathway on which the test compound acts may be determined.

The Example below provides one method for performing such assays.

EXAMPLE 10

Identification of the Pathway in which a Proliferation-Required Gene Lies or the Pathway on which an Antibiotic Acts

A. Preparation of Bacterial Stocks for Assay

To provide a consistent source of cells to screen, frozen stocks of host bacteria containing the desired antisense construct are prepared using standard microbiological techniques. For example, a single clone of the organism can be isolated by streaking out a sample of the original stock onto an agar plate containing nutrients for cell growth and an antibiotic for which the antisense construct contains a gene which confers resistance. After overnight growth an isolated

colony is picked from the plate with a sterile needle and transferred to an appropriate liquid growth media containing the antibiotic required for maintenance of the plasmid. The cells are incubated at 30°C to 37°C with vigorous shaking for 4 to 6 hours to yield a culture in exponential growth. Sterile glycerol is added to 15% (volume to volume) and 100μL to 500 μL aliquots are distributed into sterile cryotubes, snap frozen in liquid nitrogen, and stored at -80°C for future assays.

B. Growth of Bacteria for Use in the Assay

A day prior to an assay, a stock vial is removed from the freezer, rapidly thawed (37°C water bath) and a loop of culture is streaked out on an agar plate containing nutrients for cell growth and an antibiotic to which the antisense construct confers resistance. After overnight growth at 37°C, ten randomly chosen, isolated colonies are transferred from the plate (sterile inoculum loop) to a sterile tube containing 5 mL of LB medium containing the antibiotic to which the antisense vector confers resistance. After vigorous mixing to form a homogeneous cell suspension, the optical density of the suspension is measured at 600 nm (OD_{600}) and if necessary an aliquot of the suspension is diluted into a second tube of 5 mL, sterile, LB medium plus antibiotic to achieve an $OD_{600} \leq 0.02$ absorbance units. The culture is then incubated at 37° C for 1-2 hrs with shaking until the OD_{600} reaches OD 0.2 – 0.3. At this point the cells are ready to be used in the assay.

C. Selection of Media to be Used in Assay

Two fold dilution series of the inducer are generated in culture media containing the appropriate antibiotic for maintenance of the antisense construct. Several media are tested side by side and three to four wells are used to evaluate the effects of the inducer at each concentration in each media. For example, M9 minimal media, LB broth, TBD broth and Muller-Hinton media may be tested with the inducer IPTG at the following concentrations, 50 μM, 100 μM, 200 μM, 400 μM, 600 μM, 800 μM and 1000 μM. Equal volumes of test media-inducer and cells are added to the wells of a 384 well microtiter plate and mixed. The cells are prepared as described above and diluted 1:100 in the appropriate media containing the test antibiotic immediately prior to addition to the microtiter plate wells. For a control, cells are also added to several wells of each media that do not contain inducer, for example 0 μM IPTG. Cell growth is monitored continuously by incubation at 37°C in a microtiter plate reader monitoring the OD_{600} of the wells over an 18-hour period. The percent inhibition of growth produced by each concentration of inducer is calculated by comparing the rates of logarithmic growth against that exhibited by cells growing in media without inducer. The medium yielding greatest sensitivity to inducer is selected for use in the assays described below.

D. Measurement of Test Antibiotic Sensitivity in the Absence of Antisense Construct Induction

Two-fold dilution series of antibiotics with a known mechanism of action are generated in the culture media selected for further assay development that has been supplemented with the antibiotic used to maintain the construct. A panel of test antibiotics known to act on different

pathways is tested side by side with three to four wells being used to evaluate the effect of a test antibiotic on cell growth at each concentration. Equal volumes of test antibiotic and cells are added to the wells of a 384 well microtiter plate and mixed. Cells are prepared as described above using the media selected for assay development supplemented with the antibiotic required to maintain the antisense construct and are diluted 1:100 in identical media immediately prior to addition to the microtiter plate wells. For a control, cells are also added to several wells that contain the solvent used to dissolve the antibiotics but no antibiotic. Cell growth is monitored continuously by incubation at 37°C in a microtiter plate reader monitoring the OD_{600} of the wells over an 18-hour period. The percent inhibition of growth produced by each concentration of antibiotic is calculated by comparing the rates of logarithmic growth against that exhibited by cells growing in media without antibiotic. A plot of percent inhibition against log[antibiotic concentration] allows extrapolation of an IC_{50} value for each antibiotic.

E. Measurement of Test Antibiotic Sensitivity in the Presence of Antisense Construct Inducer

The culture media selected for use in the assay is supplemented with inducer at concentrations shown to inhibit cell growth by 50% and 80% as described above and the antibiotic used to maintain the construct. Two fold dilution series of the panel of test antibiotics used above are generated in each of these media. Several antibiotics are tested side by side with three to four wells being used to evaluate the effects of an antibiotic on cell growth at each concentration, in each media. Equal volumes of test antibiotic and cells are added to the wells of a 384 well microtiter plate and mixed. Cells are prepared as described above using the media selected for use in the assay supplemented with the antibiotic required to maintain the antisense construct. The cells are diluted 1:100 into two 50 mL aliquots of identical media containing concentrations of inducer that have been shown to inhibit cell growth by 50% and 80 % respectively and incubated at 37°C with shaking for 2.5 hours. Immediately prior to addition to the microtiter plate wells, the cultures are adjusted to an appropriate OD_{600} (typically 0.002) by dilution into warm (37°C) sterile media supplemented with identical concentrations of the inducer and antibiotic used to maintain the antisense construct. For a control, cells are also added to several wells that contain solvent used to dissolve test antibiotics but which contain no antibiotic. Cell growth is monitored continuously by incubation at 37°C in a microtiter plate reader monitoring the OD_{600} of the wells over an 18-hour period. The percent inhibition of growth produced by each concentration of antibiotic is calculated by comparing the rates of logarithmic growth against that exhibited by cells growing in media without antibiotic. A plot of percent inhibition against log[antibiotic concentration] allows extrapolation of an IC_{50} value for each antibiotic.

F. Determining the Specificity of the Test Antibiotics

A comparison of the IC_{50}s generated by antibiotics of known mechanism of action under antisense induced and non-induced conditions allows the pathway in which a proliferation-required

nucleic acid lies to be identified. If cells expressing an antisense nucleic acid against a proliferation-required gene are selectively sensitive to an antibiotic acting via a particular pathway, then the gene against which the antisense acts is involved in the pathway in which the antibiotic acts.

G. Identification of Pathway in which a Test Antibiotic Acts

As discussed above, the cell-based assay may also be used to determine the pathway against which a test antibiotic acts. In such an analysis, the pathways against which each member of a panel of antisense nucleic acids acts are identified as described above. A panel of cells, each containing an inducible nucleic acid complementary to a gene in a known proliferation-required pathway, is contacted with a test antibiotic for which it is desired to determine the pathway on which it acts under inducing an non-inducing conditions. If heightened sensitivity is observed in induced cells expressing antisense complementary to a gene in a particular pathway but not in induced cells expressing antisense complementary to genes in other pathways, then the test antibiotic acts against the pathway for which heightened sensitivity was observed.

One skilled in the art will appreciate that further optimization of the assay conditions, such as the concentration of inducer used to induce antisense expression and/or the growth conditions used for the assay (for example incubation temperature and media components) may further increase the selectivity and/or magnitude of the antibiotic sensitization exhibited.

The following example confirms the effectiveness of the methods described above.

EXAMPLE 11

Identification of the Pathway in which a Proliferation-Required Gene Lies

Antibiotics of various chemical classes and modes of action were purchased from Sigma Chemicals (St. Louis, MO). Stock solutions were prepared by dissolving each antibiotic in an appropriate aqueous solution based on information provided by the manufacturer. The final working solution of each antibiotic contained no more than 0.2% (w/v) of any organic solvent. To determine their potency against a bacterial strain engineered for expression of an antisense complementary to a proliferation-required gene encoding 50S ribosomal protein, each antibiotic was serially diluted two or three fold in growth medium supplemented with the appropriate antibiotic for maintenance of the anti-sense construct. At least ten dilutions were prepared for each antibiotic. 25 µL aliquots of each dilution were transferred to discrete wells of a 384-well microplate (the assay plate) using a multi-channel pipette. Quadruplicate wells were used for each dilution of an antibiotic under each treatment condition (plus and minus inducer). Each assay plate contained twenty wells for cell growth controls (growth media replacing antibiotic), ten wells for each treatment (plus and minus inducer, in this example IPTG). Assay plates were usually divided into the two treatments: half the plate containing induced cells and an appropriate concentrations of inducer (in this example IPTG) to maintain the state of induction, the other half containing non-induced cells in the absence of IPTG.

Cells for the assay were prepared as follows. Bacterial cells containing a construct, from which expression of antisense nucleic acid complementary to *rplL* and *rplJ*, which encode proliferation-required 50S ribosomal subunit proteins, is inducible in the presence of IPTG, were grown into exponential growth (OD_{600} 0.2 to 0.3) and then diluted 1:100 into fresh media containing either 400 μM or 0 μM inducer (IPTG). These cultures were incubated at 37° C for 2.5 hr. After a 2.5 hr incubation, induced and non-induced cells were respectively diluted into an assay medium at a final OD_{600} value of 0.0004. The medium contained an appropriate concentration of the antibiotic for the maintenance of the anti-sense construct. In addition, the medium used to dilute induced cells was supplemented with 800 μM IPTG so that addition to the assay plate would result in a final IPTG concentration of 400 μM. Induced and non-induced cell suspensions were dispensed (25 μl/well) into the appropriate wells of the assay plate as discussed previously. The plate was then loaded into a plate reader, incubated at constant temperature, and cell growth was monitored in each well by the measurement of light scattering at 595 nm. Growth was monitored every 5 minutes until the cell culture attained a stationary growth phase. For each concentration of antibiotic, a percentage inhibition of growth was calculated at the time point corresponding to mid-exponential growth for the associated control wells (no antibiotic, plus or minus IPTG). For each antibiotic and condition (plus or minus IPTG), a plot of percent inhibition versus log of antibiotic concentration was generated and the IC_{50} determined. A comparison of the IC_{50} for each antibiotic in the presence and absence of IPTG revealed whether induction of the antisense construct sensitized the cell to the mechanism of action exhibited by the antibiotic. Cells which exhibited a significant (standard statistical analysis) numerical decrease in the IC_{50} value in the presence of inducer were considered to have an increased sensitivity to the test antibiotic.

The results are provided in the table below, which lists the classes and names of the antibiotics used in the analysis, the targets of the antibiotics, the IC_{50} in the absence of IPTG, the IC_{50} in the presence of IPTG, the concentration units for the IC_{50}s, the fold increase in IC_{50} in the presence of IPTG, and whether increased sensitivity was observed in the presence of IPTG.

TABLE IV
Effect of Expression of Antisense RNA to rplL and rplJ on Antibiotic Sensitivity

ANTIBIOTIC CLASS /Names	TARGET	IC50 (-IPTG)	IC50 (+IPTG)	Conc. Unit	Fold Increase in Sensitivity	Sensitivity Increased?
PROTEIN SYNTHESIS INHIBITOR ANTIBIOTICS						
AMINOGLYCOSIDES						
Gentamicin	30S ribosome function	2715	19.19	ng/ml	141	Yes
Streptomycin	30S ribosome function	11280	161	ng/ml	70	Yes
Spectinomycin	30S ribosome function	18050	<156	ng/ml		Yes
Tobramycin	30S ribosome function	3594	70.58	ng/ml	51	Yes
MACROLIDES						
Erythromycin	50S ribosome function	7467	187	ng/ml	40	Yes
AROMATIC POYKETIDES						
Tetracycline	30S ribosome function	199.7	1.83	ng/ml	109	Yes
Minocycline	30S ribosome function	668.4	3.897	ng/ml	172	Yes
Doxycycline	30S ribosome function	413.1	27.81	ng/ml	15	Yes
OTHER PROTEIN SYNTHESIS INHIBITORS						
Fusidic acid	Elongation Factor G function	59990	641	ng/ml	94	Yes
Chloramphenicol	30S ribosome function	465.4	1.516	ng/ml	307	Yes
Lincomycin	50S ribosome function	47150	324.2	ng/ml	145	Yes

ANTIBIOTIC CLASS /Names	TARGET	IC50 (-IPTG)	IC50 (+IPTG)	Conc. Unit	Fold Increase in Sensitivity	Sensitivity Increased?
OTHER ANTIBIOTIC MECHANISMS						
B-LACTAMS						
Cefoxitin	Cell wall biosynthesis	2782	2484	ng/ml	1	No
Cefotaxime	Cell wall biosynthesis	24.3	24.16	ng/ml	1	No
DNA SYNTHESIS INHIBITORS						
Nalidixic acid	DNA Gyrase activity	6973	6025	ng/ml	1	No
Ofloxacin	DNA Gyrase activity	49.61	45.89	ng/ml	1	No
OTHER						
Bacitracin	Cell membrane function	4077	4677	mg/ml	1	No
Trimethoprim	Dihydrofolate Reductase activity	128.9	181.97	ng/ml	1	No
Vancomycin	Cell wall biosynthesis	145400	72550	ng/ml	2	No

The above results demonstrate that induction of an antisense RNA to genes encoding 50S ribosomal subunit proteins results in a selective and highly significant sensitization of cells to antibiotics that inhibit ribosomal function and protein synthesis. The above results further demonstrate that induction of an antisense construct to an essential gene sensitizes an organism to compounds that interfere with that gene products' biological role. This sensitization is restricted to compounds that interfere with pathways associated with the targeted gene and it's product.

Assays utilizing antisense constructs to essential genes can be used to identify compounds that specifically interfere with the activity of multiple targets in a pathway. Such constructs can be used to simultaneously screen a sample against multiple targets in one pathway in one reaction (Combinatorial HTS).

Furthermore, as discussed above, panels of antisense construct containing cells may be used to characterize the point of intervention of any compound affecting an essential biological pathway including antibiotics with no known mechanism of action.

Another embodiment of the present invention is a method for determining the pathway against which a test antibiotic compound is active in which the activity of target proteins or nucleic acids involved in proliferation-required pathways is reduced by contacting cells with a sub-lethal concentration of a known antibiotic which acts against the target protein or nucleic acid. In one embodiment, the target protein or nucleic acid is a target protein or nucleic acid corresponding to a proliferation-required nucleic acid identified using the methods described above. The method is similar to those described above for determining which pathway a test antibiotic acts against except that rather than reducing the activity or level of a proliferation-required gene product using a sub-lethal level of antisense to a proliferation-required nucleic acid, the activity or level of the proliferation-required gene product is reduced using sub-lethal level of a known antibiotic which acts against the proliferation required gene product.

Interactions between drugs which affect the same biological pathway has been described in the literature. For example, Mecillinam (Amdinocillin) binds to and inactivates the penicillin binding protein 2 (PBP2, product of the *mrdA* in *E. coli*). This antibiotic inteacts with other antibiotics that inhibit PBP2 as well as antibiotics that inhibit other penicillin binding proteins such as PBP3 [(Gutmann, L., Vincent, S., Billot-Klein, D., Acar, J.F., Mrena, E., and Williamson, R. (1986) Involvement of penicillin-binding protein 2 with other penicillin-binding proteins in lysis of *Escherichia coli* by some beta-lactam antibiotics alone and in synergistic lytic effect of amdinocillin (mecillinam). Antimicrobial Agents & Chemotherapy, 30:906-912)]. Interactions between drugs could, therefore, involve two drugs that inhibit the same target protein or nucleic acid or inhibit different proteins or nucleic acids in the same pathway [(Fukuoka, T., Domon, H., Kakuta, M., Ishii, C., Hirasawa, A., Utsui, Y., Ohya, S., and Yasuda, H. (1997) Combination effect between panipenem and vancomycin on highly methicillin-resistant Staphylococcus aureus. Japan. J. Antibio. 50:411-419; Smith, C.E., Foleno, B.E., Barrett, J.F., and Frosc, M.B. (1997) Assessment

of the synergistic interactions of levofloxacin and ampicillin against *Enterococcus faecium* by the checkerboard agar dilution and time-kill methods. Diagnos. Microbiol. Infect. Disease 27:85-92; den Hollander, J.G., Horrevorts, A.M., van Goor, M.L., Verbrugh, H.A., and Mouton, J.W. (1997) Synergism between tobramycin and ceftazidime against a resistant Pseudomonas aeruginosa strain, tested in an in vitro pharmacokinetic model. Antimicrobial Agents & Chemotherapy. 41:95-110)].

Two drugs may interact even though they inhibit different targets. For example, the proton pump inhibitor, Omeprazole, and the antibiotic, Amoxycillin, two synergistic compounds acting together, can cure *Helicobacter pylori* infection [(Gabryelewicz, A., Laszewicz, W., Dzieniszewski, J., Ciok, J., Marlicz, K., Bielecki, D., Popiela, T., Legutko, J., Knapik, Z., Poniewierka, E. (1997) Multicenter evaluation of dual-therapy (omeprazol and amoxycillin) for *Helicobacter pylori*-associated duodenal and gastric ulcer (two years of the observation). J. Physiol. Pharmacol. 48 Suppl 4:93-105)].

The growth inhibition from the sub-lethal concentration of the known antibiotic may be at least about 5%, at least about 8%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, or at least about 75%, or more.

Alternatively, the sub-lethal concentration of the known antibiotic may be determined by measuring the activity of the target proliferation-required gene product rather than by measuring growth inhibition.

Cells are contacted with a combination of each member of a panel of known antibiotics at a sub-lethal level and varying concentrations of the test antibiotic. As a control, the cells are contacted with varying concentrations of the test antibiotic alone. The IC_{50} of the test antibiotic in the presence and absence of the known antibiotic is determined. If the IC_{50}s in the presence and absence of the known drug are substantially similar, then the test drug and the known drug act on different pathways. If the IC_{50}s are substantially different, then the test drug and the known drug act on the same pathway.

Another embodiment of the present invention is a method for identifying a candidate compound for use as an antibiotic in which the activity of target proteins or nucleic acids involved in proliferation-required pathways is reduced by contacting cells with a sub-lethal concentration of a known antibiotic which acts against the target protein or nucleic acid. In one embodiment, the target protein or nucleic acid is a target protein or nucleic acid corresponding to a proliferation-required nucleic acid identified using the methods described above. The method is similar to those described above for identifying candidate compounds for use as antibiotics except that rather than reducing the activity or level of a proliferation-required gene product using a sub-lethal level of antisense to a proliferation-required nucleic acid, the activity or level of the proliferation-required gene product is reduced using a sub-lethal level of a known antibiotic which acts against the proliferation required gene product.

The growth inhibition from the sub-lethal concentration of the known antibiotic may be at least about 5%, at least about 8%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, or at least about 75%, or more.

Alternatively, the sub-lethal concentration of the known antibiotic may be determined by measuring the activity of the target proliferation-required gene product rather than by measuring growth inhibition.

In order to characterize test compounds of interest, cells are contacted with a panel of known antibiotics at a sub-lethal level and one or more concentrations of the test compound. As a control, the cells are contacted with the same concentrations of the test compound alone. The IC_{50} of the test compound in the presence and absence of the known antibiotic is determined. If the IC_{50} of the test compound is substantially different in the presence and absence of the known drug then the test compound is a good candidate for use as an antibiotic. As discussed above, once a candidate compound is identified using the above methods its structure may be optimized using standard techniques such as combinatorial chemistry.

Representative known antibiotics which may be used in each of the above methods are provided in the table below. However, it will be appreciated that other antibiotics may also be used.

ANTIBIOTIC	INHIBITS/TARGET	RESISTANT MUTANTS
Inhibitors of Transcription		
Rifamycin, 1959 Rifampicin Rifabutin Rifaximin	Inhibits initiation of transcription/ß-subunit RNA polymerase, *rpoB*	*rpoB, crp, cyaA*
Streptolydigin	Accelerates transcription chain termination/ß-subunit RNA polymerase	*rpoB*
Streptovaricin	an acyclic ansamycin, inhibits RNA polymerase	*rpoB*
Actinomycin D+EDTA	Intercalates between 2 successive G-C pairs, *rpoB*, inhibits RNA synthesis	*pldA*
Inhibitors of Nucleic Acid Metabolism		
Quinolones, 1962	α subunit gyrase and/or topoisomerase IV,	
Nalidixic acid Oxolinic acid	*gyrA*	*gyrAorB, icd, sloB*
Fluoroquinolones Ciprofloxacin, 1983 Norfloxacin	α subunit gyrase, *gyrA* and/or topoisomerase IV (probable target in Staph)	*gyrA* *norA* (efflux in Staph) *hipQ*
Coumerins	Inhibits ATPase activity of ß-subunit	
Novobiocin	gyrase, *gyrB*	*gyrB, cysB, cysE, nov, ompA*
Coumermycin	Inhibits ATPase activity of ß-subunit gyrase, *gyrB*	*gyrB, hisW*
Albicidin	DNA synthesis	*tsx* (nucleoside channel)
Metronidazole	Causes single-strand breaks in DNA	*nar*

Inhibitors of Metabolic Pathways

Sulfonamides, 1932 Sulfanilamide	blocks synthesis of dihydrofolate,dihydro-pteroate synthesis, *folP*	*folP, gpt, pabA, pabB, pabC*
Trimethoprim, 1962	Inhibits dihydrofolate reductase, *folA*	*folA, thyA*
Showdomycin	Nucleoside analogue capable of alkylating sulfhydryl groups, inhibitor of thymidylate synthetase	*nupC, pnp*
Thiolactomycin	type II fatty acid synthase inhibitor	*emrB* *fadB, emrB* due to gene dosage
Psicofuranine	Adenosine glycoside antibiotic, target is GMP synthetase	*guaA,B*
Triclosan	Inhibits fatty acid synthesis	*fabI (envM)*
Diazoborines Isoniazid Ethionamide	heterocyclic, contains boron, inhibit fatty acid synthesis, enoyl-ACP reductase, *fabI*	*fabI (envM)*

Inhibitors of Translation

Phenylpropanoids Chloramphenicol, 1947	Binds to ribosomal peptidyl transfer center preventing peptide translocation/ binds to S6, L3, L6, L14, L16, L25, L26, L27, but preferentially to L16	*rrn, cmlA, marA, ompF, ompR*
Tetracyclines, 1948, type II polyketides Minocycline Doxycycline	Binding to 30S ribosomal subunit, "A" site on 30S subunit, blocks peptide elongation, strongest binding to S7	*clmA (cmr), mar, ompF*
Macrolides (type I polyketides) Erythromycin, 1950 Carbomycin Spiramycin, etc	Binding to 50 S ribosomal subunit, 23S rRNA, blocks peptide translocation, L15, L4, L12	*rrn, rplC, rplD, rplV, mac*
Aminoglycosides	Irreversible binding to 30S ribosomal subunit, prevents translation or causes mistranslation of mRNA/16S rRNA	
Streptomycin, 1944		*rpsL, strC,M, ubiF*
Neomycin		*atpA-E, ecfB, hemAC,D,E,G, topA,*
Spectinomycin		*rpsC,D,E, rrn, spcB*
Kanamycin		*atpA-atpE, cpxA, ecfB, hemA,B,L, topA*
Kasugamycin		*ksgA,B,C,D, rplB,K, rpsI,N,M,R*
Gentamicin, 1963		*rplF, ubiF*
Amikacin		*cpxA*
Paromycin		*rpsL*
Lincosamides Lincomycin, 1955 Clindamycin	Binding to 50 S ribosomal subunit, blocks peptide translocation	*linB, rplN,O, rpsG*
Streptogramins Virginiamycin, 1955 Pristinamycin Synercid: quinupristin /dalfopristin	2 components, Streptogramins A&B, bind to the 50S ribosomal subunit blocking peptide translocation and peptide bond formation	
Fusidanes Fusidic Acid	Inhibition of elongation factor G (EF-G) prevents peptide translocation	*fusA*
Kirromycin (Mocimycin)	Inhibition of elongation factor TU (EF-Tu), prevents peptide bond formation	*tufA,B*
Pulvomycin	Binds to and inhibits EF-TU	
Thiopeptin	Sulfur-containing antibiotic, inhibits protein synthesis,EF-G	*rplE*
Tiamulin	Inhibits protein synthesis	*rplC, rplD*

Negamycin	Inhibits termination process of protein synthesis	*prfB*
Oxazolidinones Linezolid	23S rRNA	
Isoniazid		*pdx*
Nitrofurantoin	Inhibits protein synthesis, nitroreductases convert nitrofurantoin to highly reactive electrophilic intermediates which attack bacterial ribosomal proteins non-specifically	*nfnA,B*
Pseudomonic Acids Mupirocin (Bactroban)	Inhibition of isoleucyl tRNA synthetase-used for Staph, topical cream, nasal spray	*ileS*
Indolmycin	Inhibits tryptophanyl-tRNA synthetase	*trpS*
Viomycin		*rrmA* (23S rRNA methyltransferase; mutant has slow growth rate, slow chain elongation rate, and viomycin resistance)
Thiopeptides	Binds to L11-23S RNA complex	
Thiostrepton	Inhibits GTP hydrolysis by EF-G	
Micrococcin	Stimulates GTP hydrolysis by EF-G	

Inhibitors of Cell Walls/Membranes

ß-lactams Penicillin, 1929 Ampicillin Methicillin, 1960	Inhibition of one or more cell wall transpeptidases, endopeptidases, and glycosidases (PBPs), of the 12 PBPs only 2 are essential: *mrdA* (PBP2) and *ftsI (pbpB,* PBP3)	*ampC, ampD, ampE, envZ, galU, hipA, hipQ, ompC, ompF, ompR, ptsI, rfa, tolD, tolE*
Cephalosporins, 1962		*tonB*
Mecillinam (amdinocillin)	Binds to and inactivates PBP2 (*mrdA*)	*alaS, argS, crp, cyaA, envB, mrdA,B, mreB,C,D*
Aztreonam (Furazlocillin)	Inactivates PBP3 (*ftsI*)	
Bacilysin, Tetaine	Dipeptide, inhib glucosamine synthase	*dppA*
Glycopeptides Vancomycin, 1955	Inhib G+ cell wall syn, binds to terminal D-ala-D-ala of pentapeptide,	
Polypeptides Bacitracin	Prevents dephosphorylation and regeneration of lipid carrier	*rfa*
Cyclic lipopeptide Daptomycin, 1980	Disrupts multiple aspects of membrane function, including peptidoglycan synthesis, lipoteichoic acid synthesis, and the bacterial membrane potential	
Cyclic polypeptides Polymixin, 1939	Surfactant action disrupts cell membrane lipids, binds lipid A mioety of LPS	*pmrA*
Fosfomycin, 1969	Analogue of P-enolpyruvate, inhibits 1st step in peptidoglycan synthesis - UDP-N-acetylglucosamine enolpyruvyl transferase, *murA*. Also acts as Immunosuppressant	*murA, crp, cyaA glpT, hipA, ptsI, uhpT*
Cycloserine	Prevents formation of D-ala dimer, inhibits D-ala ligase, *ddlA,B*	*hipA, cycA*

Alafosfalin	phosphonodipeptide, cell wall synthesis inhibitor, potentiator of β-lactams	*pepA, tpp*
Inhibitors of Protein Processing/Transport		
Globomycin	Inhibits signal peptidase II (cleaves prolipoproteins subsequent to lipid modification, *lspA*	*lpp, dnaE*

EXAMPLE 12

Transfer of Exogenous Nucleic Acid Sequences to other Bacterial Species Using the *E. coli* Expression Vectors or Expression Vectors Functional in Bacterial Species other than *E. coli.*

Molecule No. EcXA190, encoding a portion of the *b3052* gene of *Escherichia coli*, was either transformed directly into *Enterobacter cloacae, Salmonella typhimurium* and/or *Klebsiella pneumoniae* or subcloned into an expression vector functional in these species and the subclones transformed into these organisms. Suitable expression vectors are well known in the art. These expression vectors were introduced into *Enterobacter cloacae, Salmonella typhimurium* and/or *Klebsiella pneumoniae* cells that were then assayed for growth inhibition according to the method of Example 1. After growth in liquid culture, cells were plated at various serial dilutions and a score determined by calculating the log difference in growth for INDUCED vs. UNINDUCED antisense RNA expression as determined by the maximum 10 fold dilution at which a colony was observed. If there was no effect of antisense RNA expression in one organism, the clone is given a score of zero "0" in that organism. In contrast, a score of "8" means that 10^8 times more cells were required to observe a colony formed on the induced state than in the non-induced state under the conditions used and in that organism.

Expression vectors containing Molecule No. EcXA190 were found to inhibit bacterial growth in all four organisms when expression of the antisense RNA was induced with IPTG. A score of 4 was assigned for *Escherichia coli,* 6 for *Enterobacter cloacae,* and 8 for *Salmonella typhimurium* and 3 for *Klebsiella pneumoniae* (obvious additional growth defect as well). The protein encoded by this sequence may be used as a target sequence to screen candidate compound libraries as described above.

In addition, the above methods were validated using other antisense nucleic acids which inhibit the growth of *E. coli* which were identified using methods similar to those described above. Expression vectors which inhibited growth of *E. coli* upon induction of antisense RNA expression with IPTG were transformed directly into *Enterobacter cloacae, Klebsiella pneumonia* or *Salmonella typhimurium.* The transformed cells were then assayed for growth inhibition according to the method of Example 1. After growth in liquid culture, cells were plated at various serial dilutions and a score determined by calculating the log difference in growth for INDUCED vs. UNINDUCED antisense RNA expression as determined by the maximum 10 fold dilution at which

a colony was observed. The results of these experiments are listed in Table V below. If there was no effect of antisense RNA expression in a microorganism, the clone is minus in Table V below. In contrast, a positive in Table V below means that at least 10 fold more cells were required to observe a colony on the induced plate than on the non-induced plate under the conditions used and in that microorganism.

TABLE V

Sensitivity of Other Microorganisms to Antisense Nucleic Acids That Inhibit Proliferation in *E. coli*

Mol. No.	*S. typhimurium*	*E. cloacae*	*K. pneumoniae*
EcXA001	+	+	-
EcXA004	+	-	-
EcXA005	+	+	+
EcXA006	-	-	-
EcXA007	-	+	-
EcXA008	+	-	+
EcXA009	-	-	-
EcXA010	+	+	+
EcXA011	-	+	-
EcXA012	-	+	-
EcXA013	+	+	+
EcXA014	+	+	-
EcXA015	+	+	+
EcXA016	+	+	+
EcXA017	+	+	+
EcXA018	+	+	+
EcXA019	+	+	+
EcXA020	+	+	+
EcXA021	+	+	+
EcXA023	+	+	+
EcXA024	+	-	+
EcXA025	-	-	-
EcXA026	+	+	-
EcXA027	+	+	-
EcXA028	+	-	-
EcXA029	-	-	-
EcXA030	+	+	+
EcXA031	+	-	-
EcXA032	+	+	-
EcXA033	+	+	+
EcXA034	+	+	+
EcXA035	-	-	-
EcXA036	+	-	+
EcXA037	+	+	-
EcXA038	+	+	+
EcXA039	+	-	-
EcXA041	+	+	+
EcXA042	-	+	+

Mol. No.	*S. typhimurium*	*E. cloacae*	*K. pneumoniae*
EcXA043	-	-	-
EcXA044	-	-	-
EcXA045	+	+	+
EcXA046	-	-	-
EcXA047	+	+	-
EcXA048	-	-	-
EcXA049	+	-	-
EcXA050	-	-	-
EcXA051	+	-	-
EcXA052	+	-	-
EcXA053	+	+	+
EcXA054	-	-	+
EcXA055	+	-	-
EcXA056	+	-	+
EcXA057	+	+	-
EcXA058	-	-	-
EcXA059	+	+	+
EcXA060	-	-	-
EcXA061	-	-	-
EcXA062	-	-	-
EcXA063	+	+	-
EcXA064	-	-	-
EcXA065	+	+	-
EcXA066	-	-	-
EcXA067	-	+	-
EcXA068	-	-	-
EcXA069	-	+	-
EcXA070	-	-	-
EcXA071	+	-	-
EcXA072	+	-	+
EcXA073	+	+	+
EcXA074	+	+	+
EcXA075	+	-	-
EcXA076	-	+	-
EcXA077	+	+	-
EcXA079	+	+	+
EcXA080	+	-	-
EcXA082	-	+	-
EcXA083	-	-	-
EcXA084	-	+	-
EcXA086	-	-	-
EcXA087	-	-	-
EcXA088	-	-	-
EcXA089	-	-	-
EcXA090	-	-	-
EcXA091	-	-	-
EcXA092	-	-	-
EcXA093	-	-	-

Mol. No.	*S. typhimurium*	*E. cloacae*	*K. pneumoniae*
EcXA094	+	+	+
EcXA095	+	+	-
EcXA096	-	-	-
EcXA097	+	-	-
EcXA098	+	-	-
EcXA099	-	-	-
EcXA100	-	-	-
EcXA101	-	-	-
EcXA102	-	-	-
EcXA103	-	+	-
EcXA104	+	+	+
EcXA106	+	+	-
EcXA107	-	-	-
EcXA108	-	-	-
EcXA109	-	-	-
EcXA110	+	+	-
EcXA111	-	-	-
EcXA112	-	+	-
EcXA113	+	+	+
EcXA114	-	+	-
EcXA115	-	+	-
EcXA116	+	+	-
EcXA117	+	-	-
EcXA118	-	-	-
EcXA119	+	+	-
EcXA120	-	-	-
EcXA121	-	-	-
EcXA122	+	-	+
EcXA123	+	-	-
EcXA124	-	-	-
EcXA125	-	-	-
EcXA126	-	-	-
EcXA127	+	+	-
EcXA128	-	-	-
EcXA129	-	+	-
EcXA130	+	+	-
EcXA132	-	-	-
EcXA133	-	-	-
EcXA136	-	-	-
EcXA137	-	-	-
EcXA138	+	-	-
EcXA139	-	-	-
EcXA140	+	-	-
EcXA141	+	-	-
EcXA142	-	-	-
EcXA143	-	+	-
EcXA144	+	+	-
EcXA145	-	-	-

Mol. No.	*S. typhimurium*	*E. cloacae*	*K. pneumoniae*
EcXA146	-	-	-
EcXA147	-	-	-
EcXA148	-	-	-
EcXA149	+	+	+
EcXA150	-	-	-
EcXA151	+	-	-
EcXA152	-	-	-
EcXA153	+	+	-
EcXA154	-	-	-
EcXA155	-	-	ND
EcXA156	-	+	-
EcXA157	-	-	-
EcXA158	-	-	-
EcXA159	+	-	-
EcXA160	+	-	-
EcXA162	-	-	-
EcXA163	-	-	-
EcXA164	-	-	-
EcXA165	-	-	-
EcXA166	-	-	-
EcXA167	-	-	-
EcXA168	-	-	-
EcXA169	-	+	-
EcXA171	-	-	-
EcXA172	-	-	-
EcXA173	-	-	-
EcXA174	-	-	-
EcXA175	-	-	-
EcXA176	-	-	-
EcXA178	-	-	-
EcXA179	-	-	-
EcXA180	+	-	-
EcXA181	-	-	-
EcXA182	-	-	-
EcXA183	-	-	-
EcXA184	-	-	-
EcXA185	-	-	-
EcXA186	-	-	-
EcXA187	+	+	+
EcXA189	+	-	-
EcXA190	+	+	+
EcXA191	+	+	-
EcXA192	-	+	-

Thus, the ability of an antisense nucleic acid which inhibits the proliferation of *E. coli* to inhibit the growth of other organims may be evaluated by either transforming the antisense nucleic acid directly into other *Escherichia* species or inserting the antisense nucleic acid into expression

vectors that are functional in other Gram negative species such as *Enterobacter cloacae, Salmonella typhimurium, and/or Klebsiella pneumoniae* . Similarly, the antisense nucleic acid can be inserted in expression vectors that are functional in Gram-positive species such as *Staphylococcus aureus, Enterococcus faecalis* and *Streptococcus pneumoniae* or other species.

Those skilled in the art will appreciate that a negative result in a heterologous microorganism does not mean that that microorganism is missing that gene nor does it mean that the gene is unessential. However, a positive result means that the heterologous microorganism contains a homologous gene which is required for proliferation of that microorganism. The homologous gene may be obtained using the methods described herein. Those cells that are inhibited by antisense may be used in cell-based assays as described herein for the identification and characterization of compounds in order to develop antibiotics effective in these microorganisms. Those skilled in the art will appreciate that an antisense molecule which works in the microorganism from which it was obtained will not always work in a heterologous microorganism.

EXAMPLE 13

Use of Identified Exogenous Nucleic Acid Sequences as Probes

The identified sequence of the present invention can be used as probes to obtain the sequence of additional genes of interest from a second organism. For example, probes to genes encoding potential bacterial target proteins may be hybridized to nucleic acids from other organisms including other bacteria and higher organisms, to identify homologous sequences. Such hybridization might indicate that the protein encoded by the gene to which the probe corresponds is found in humans and therefore not necessarily a good drug target. Alternatively, the gene can be conserved only in bacteria and therefore would be a good drug target for a broad spectrum antibiotic or antimicrobial.

Probes derived from the identified nucleic acid sequences of interest or portions thereof can be labeled with detectable labels familiar to those skilled in the art, including radioisotopes and non-radioactive labels, to provide a detectable probe. The detectable probe can be single stranded or double stranded and can be made using techniques known in the art, including *in vitro* transcription, nick translation, or kinase reactions. A nucleic acid sample containing a sequence capable of hybridizing to the labeled probe is contacted with the labeled probe. If the nucleic acid in the sample is double stranded, it can be denatured prior to contacting the probe. In some applications, the nucleic acid sample can be immobilized on a surface such as a nitrocellulose or nylon membrane. The nucleic acid sample can comprise nucleic acids obtained from a variety of sources, including genomic DNA, cDNA libraries, RNA, or tissue samples.

Procedures used to detect the presence of nucleic acids capable of hybridizing to the detectable probe include well known techniques such as Southern blotting, Northern blotting, dot blotting, colony hybridization, and plaque hybridization. In some applications, the nucleic acid capable of hybridizing to the labeled probe can be cloned into vectors such as expression vectors, sequencing vectors, or in

vitro transcription vectors to facilitate the characterization and expression of the hybridizing nucleic acids in the sample. For example, such techniques can be used to isolate, purify and clone sequences from a genomic library, made from a variety of bacterial species, which are capable of hybridizing to probes made from the sequences identified in Examples 5 and 6.

EXAMPLE 14

Preparation of PCR Primers and Amplification of DNA

The identified *E. coli* genes corresponding directly to or located within the operon of nucleic acid sequences required for proliferation or portions thereof can be used to prepare PCR primers for a variety of applications, including the identification or isolation of homologous sequences from other species, for example *S. typhimurium, E. cloacae, E. faecalis, S. pneumoniae, and K. pneumoniae,* which contain part or all of the homologous genes. Because homologous genes are related but not identical in sequence, those skilled in the art will often employ degenerate sequence PCR primers. Such degenerate sequence primers are designed based on conserved sequence regions, either known or suspected, such as conserved coding regions. The successful production of a PCR product using degenerate probes generated from the sequences identified herein would indicate the presence of a homologous gene sequence in the species being screened. The PCR primers are at least 10 nucleotides, and preferably at least 20 nucleotides in length. More preferably, the PCR primers are at least 20-30 nucleotides in length. In some embodiments, the PCR primers can be more than 30 nucleotides in length. It is preferred that the primer pairs have approximately the same G/C ratio, so that melting temperatures are approximately the same. A variety of PCR techniques are familiar to those skilled in the art. For a review of PCR technology, see Molecular Cloning to Genetic Engineering White, B.A. Ed. in Methods in Molecular Biology 67: Humana Press, Totowa 1997. When the entire coding sequence of the target gene is known, the 5' and 3' regions of the target gene can be used as the sequence source for PCR probe generation. In each of these PCR procedures, PCR primers on either side of the nucleic acid sequences to be amplified are added to a suitably prepared nucleic acid sample along with dNTPs and a thermostable polymerase such as Taq polymerase, Pfu polymerase, or Vent polymerase. The nucleic acid in the sample is denatured and the PCR primers are specifically hybridized to complementary nucleic acid sequences in the sample. The hybridized primers are extended. Thereafter, another cycle of denaturation, hybridization, and extension is initiated. The cycles are repeated multiple times to produce an amplified fragment containing the nucleic acid sequence between the primer sites.

EXAMPLE 15

Inverse PCR

The technique of inverse polymerase chain reaction can be used to extend the known nucleic acid sequence identified in Examples 5 and 6. The inverse PCR reaction is described generally by Ochman et al., in Ch. 10 of PCR Technology: Principles and Applications for DNA Amplification, (Henry A. Erlich, Ed.) W.H. Freeman and Co. (1992). Traditional PCR requires two primers that are

used to prime the synthesis of complementary strands of DNA. In inverse PCR, only a core sequence need be known.

Using the sequences identified as relevant from the techniques taught in Examples 5 and 6 and applied to other species of bacteria, a subset of exogenous nucleic sequences are identified that correspond to genes or operons that are required for bacterial proliferation. In species for which a genome sequence is not known, the technique of inverse PCR provides a method for obtaining the gene in order to determine the sequence or to place the probe sequences in full context to the target sequence to which the identified exogenous nucleic acid sequence binds.

To practice this technique, the genome of the target organism is digested with an appropriate restriction enzyme so as to create fragments of nucleic acid that contain the identified sequence as well as unknown sequences that flank the identified sequence. These fragments are then circularized and become the template for the PCR reaction. PCR primers are designed in accordance with the teachings of Example 15 and directed to the ends of the identified sequence. The primers direct nucleic acid synthesis away from the known sequence and toward the unknown sequence contained within the circularized template. After the PCR reaction is complete, the resulting PCR products can be sequenced so as to extend the sequence of the identified gene past the core sequence of the identified exogenous nucleic acid sequence identified. This process can be repeated iteratively if necessary. In this manner, the full sequence of each novel gene can be identified. Additionally the sequences of adjacent coding and noncoding regions can be identified.

EXAMPLE 16

Identification of Genes Required for *Staphylococcus aureus* Proliferation

Genes required for proliferation in *Staphylococcus aureus* are identified according to the methods described above.

EXAMPLE 17

Identification of Genes Required for *Neisseria gonorrhoeae* Proliferation

Genes required for proliferation in *Neisseria gonorrhoeae* are identified according to the methods described above.

EXAMPLE 18

Identification of Genes Required for *Pseudomonas aeruginosa* Proliferation

Genes required for proliferation in *Pseudomonas aeruginosa* are identified according to the methods described above.

EXAMPLE 19

Identification of Genes Required for *Enterococcus faecalis* Proliferation

Genes required for proliferation in *Enterococcus faecalis* are identified according to the methods described above.

EXAMPLE 20

Identification of Genes Required for *Haemophilus influenzae* Proliferation

Genes required for proliferation in *Haemophilus influenzae* are identified according to the methods described above.

EXAMPLE 21

Identification of Genes Required for *Salmonella typhimurium* Proliferation

Genes required for proliferation in *Salmonella typhimurium* are identified according to the methods described above.

EXAMPLE 22

Identification of Genes Required for *Helicobacter pylori* Proliferation

Genes required for proliferation in *Helicobacter pylori* are identified according to the methods described above.

EXAMPLE 23

Identification of Genes Required for *Mycoplasma pneumoniae* Proliferation

Genes required for proliferation in *Mycoplasma pneumoniae* are identified according to the methods described above.

EXAMPLE 24

Identification of Genes Required for *Plasmodium ovale* Proliferation

Genes required for proliferation in *Plasmodium ovale* are identified according to the methods described above.

EXAMPLE 25

Identification of Genes Required for *Saccharomyces cerevisiae* Proliferation

Genes required for proliferation in *Saccharomyces cerevisiae* are identified according to the methods described above.

EXAMPLE 26

Identification of Genes Required for *Entamoeba histolytica* Proliferation

Genes required for proliferation in *Entamoeba histolytica* are identified according to the methods described above.

EXAMPLE 27

Identification of Genes Required for *Candida albicans* Proliferation

Genes required for proliferation in *Candida albicans* are identified according to the methods described above.

EXAMPLE 28

Identification of Genes Required for *Klebsiella pneumoniae* Proliferation

Genes required for proliferation in *Klebsiella pneumoniae* are identified according to the methods described above.

EXAMPLE 29

Identification of Genes Required for *Salmonella typhi* Proliferation

Genes required for proliferation in *Salmonella typhi* are identified according to the methods described above.

EXAMPLE 30

Identification of Genes Required for *Salmonella paratyphi* Proliferation

Genes required for proliferation in *Salmonella paratyphi* are identified according to the methods described above.

EXAMPLE 31

Identification of Genes Required for *Salmonella cholerasuis* Proliferation

Genes required for proliferation in *Salmonella cholerasuis* are identified according to the methods described above.

EXAMPLE 32

Identification of Genes Required for *Staphylococcus epidermis* Proliferation

Genes required for proliferation in *Staphylococcus epidermis* are identified according to the methods described above.

EXAMPLE 33

Identification of Genes Required for *Mycobacterium tuberculosis* Proliferation

Genes required for proliferation in *Mycobacterium tuberculosis* are identified according to the methods described above.

EXAMPLE 34

Identification of Genes Required for *Mycobacterium leprae* Proliferation

Genes required for proliferation in *Mycobacterium leprae* are identified according to the methods described above.

EXAMPLE 35

Identification of Genes Required for *Treponema pallidum* Proliferation

Genes required for proliferation in *Treponema pallidum* are identified according to the methods described above.

EXAMPLE 36

Identification of Genes Required for *Bacillus anthracis* Proliferation

Genes required for proliferation in *Bacillus anthracis* are identified according to the methods described above.

EXAMPLE 37

Identification of Genes Required for *Yersinia pestis* Proliferation

Genes required for proliferation in *Yersinia pestis* are identified according to the methods described above.

EXAMPLE 38

Identification of Genes Required for *Clostridium botulinum* Proliferation

Genes required for proliferation in *Clostridium botulinum* are identified according to the methods described above.

EXAMPLE 39

Identification of Genes Required for *Campylobacter jejuni* Proliferation

Genes required for proliferation in *Campylobacter jejuni* are identified according to the methods described above.

EXAMPLE 40

Identification of Genes Required for *Chlamydia trachomatis* Proliferation

Genes required for proliferation in *Chlamydia trachomatis* are identified according to the methods described above.

It will be appreciated that genes required for proliferation of any microorganism of interest, including those specifically mentioned herein, may be identified according to the methods described above.

Use of Isolated Exogenous Nucleic Acid Fragments as Antisense Antibiotics

In addition to using the identified sequences to enable screening of molecule libraries to identify compounds useful to identify antibiotics, the sequences themselves can be used as therapeutic agents. Specifically, the identified exogenous sequences in an antisense orientation can be provided to an individual to inhibit the translation of a bacterial target gene.

Generation of Antisense Therapeutics from Identified Exogenous Sequences

The sequences of the present invention can be used as antisense therapeutics for the treatment of bacterial infections or simply for inhibition of bacterial growth *in vitro* or *in vivo*. The therapy exploits the biological process in cells where genes are transcribed into messenger RNA (mRNA) that is then translated into proteins. Antisense RNA technology contemplates the use of antisense oligonucleotides complementary to a target gene that will bind to its target nucleic acid and decrease or inhibit the expression of the target gene. For example, the antisense nucleic acid may inhibit the translation or transcription of the target nucleic acid. In one embodiment, antisense oligonucleotides can be used to treat and control a bacterial infection of a cell culture containing a population of desired cells contaminated with bacteria. In another embodiment, the antisense oligonucleotides can be used to treat an organism with a bacterial infection.

Antisense oligonucleotides can be synthesized from any of the sequences of the present invention using methods well known in the art. In a preferred embodiment, antisense oligonucleotides are synthesized using artificial means. Uhlmann & Peymann, Chemical Rev. 90:543-584 (1990) review antisense oligonucleotide technology in detail. Modified or unmodified antisense oligonucleotides can be used as therapeutic agents. Modified antisense oligonucleotides

are preferred since it is well known that antisense oligonucleotides are extremely unstable. Modification of the phosphate backbones of the antisense oligonucleotides can be achieved by substituting the internucleotide phosphate residues with methylphosphonates, phosphorothioates, phosphoramidates, and phosphate esters. Nonphosphate internucleotide analogs such as siloxane bridges, carbonate brides, thioester bridges, as well as many others known in the art may also be used. The preparation of certain antisense oligonucleotides with modified internucleotide linkages is described in U.S. Patent No. 5,142,047.

Modifications to the nucleoside units of the antisense oligonucleotides are also contemplated. These modifications can increase the half-life and increase cellular rates of uptake for the oligonucleotides *in vivo*. For example, α-anomeric nucleotide units and modified nucleotides such as 1,2-dideoxy-d-ribofuranose, 1,2-dideoxy-1-phenylribofuranose, and N^4, N^4-ethano-5-methyl-cytosine are contemplated for use in the present invention.

An additional form of modified antisense molecules is found in peptide nucleic acids. Peptide nucleic acids (PNA) have been developed to hybridize to single and double stranded nucleic acids. PNA are nucleic acid analogs in which the entire deoxyribose-phosphate backbone has been exchanged with a chemically completely different, but structurally homologous, polyamide (peptide) backbone containing 2-aminoethyl glycine units. Unlike DNA, which is highly negatively charged, the PNA backbone is neutral. Therefore, there is much less repulsive energy between complementary strands in a PNA-DNA hybrid than in the comparable DNA-DNA hybrid, and consequently they are much more stable. PNA can hybridize to DNA in either a Watson/Crick or Hoogsteen fashion (Demidov et al., *Proc. Natl. Acad. Sci. U.S.A.* **92**:2637-2641, 1995; Egholm, *Nature* **365**:566-568, 1993; Nielsen et al., *Science* **254**:1497-1500, 1991; Dueholm et al., *New J. Chem.* **21**:19-31, 1997).

Molecules called PNA "clamps" have been synthesized which have two identical PNA sequences joined by a flexible hairpin linker containing three 8-amino-3,6-dioxaoctanoic acid units. When a PNA clamp is mixed with a complementary homopurine or homopyrimidine DNA target sequence, a PNA-DNA-PNA triplex hybrid can form which has been shown to be extremely stable (Bentin et al., *Biochemistry* **35**:8863-8869, 1996; Egholm et al., *Nucleic Acids Res.* **23**:217-222, 1995; Griffith et al., *J. Am. Chem. Soc.* **117**:831-832, 1995).

The sequence-specific and high affinity duplex and triplex binding of PNA have been extensively described (Nielsen et al., *Science* **254**:1497-1500, 1991; Egholm et al., *J. Am. Chem. Soc.* **114**:9677-9678, 1992; Egholm et al., *Nature* **365**:566-568, 1993; Almarsson et al., *Proc. Natl. Acad. Sci. U.S.A.* **90**:9542-9546, 1993; Demidov et al., *Proc. Natl. Acad. Sci. U.S.A.* **92**:2637-2641, 1995). They have also been shown to be resistant to nuclease and protease digestion (Demidov et al., *Biochem. Pharm.* **48**:1010-1313, 1994). PNA has been used to inhibit gene expression (Hanvey et al., *Science* **258**:1481-1485,1992; Nielsen et al., *Nucl. Acids. Res.*, **21**:197-200, 1993; Nielsen et al., *Gene* **149**:139-145, 1994; Good & Nielsen, **Science**, 95: 2073-2076, 1998), to block restriction enzyme activity (Nielsen et al., *supra.*, 1993), to act as an artificial transcription promoter (Mollegaard, *Proc.*

Natl. Acad. Sci. U.S.A. 91:3892-3895, 1994) and as a pseudo restriction endonuclease (Demidov et al., *Nucl. Acids. Res.* 21:2103-2107, 1993). Recently, PNA has also been shown to have antiviral and antitumoral activity mediated through an antisense mechanism (Norton, *Nature Biotechnol.*, 14:615-619, 1996; Hirschman et al., *J. Investig. Med.* 44:347-351, 1996). PNAs have been linked to various peptides in order to promote PNA entry into cells (Basu et al., *Bioconj. Chem.* 8:481-488, 1997; Pardridge et al., *Proc. Natl. Acad. Sci. U.S.A.* 92:5592-5596, 1995).

The antisense oligonucleotides contemplated by the present invention can be administered by direct application of oligonucleotides to a target using standard techniques well known in the art. The antisense oligonucleotides can be generated within the target using a plasmid, or a phage. Alternatively, the antisense nucleic acid may be expressed from a sequence in the chromosome of the target cell. For example, a promoter may be introduced into the chromosme of the target cell near the target gene such that the promoter directs teh transcription of the antisense nucleic acid. Alternatively, a nucleic acid containing the antisense sequence operably linked to a promoter may be introduced into the chromosome of the target cell. It is further contemplated that contemplated that the antisense oligonucleotide contemplated are incorporated in a ribozyme sequence to enable the antisense to specifically bind and cleave its target mRNA. For technical applications of ribozyme and antisense oligonucleotides see Rossi et al., Pharmacol. Ther. 50(2):245-254, (1991). The present invention also contemplates using a retron to introduce an antisense oligonucleotide to a cell. Retron technology is exemplified by U.S. Patent No. 5,405,775. Antisense oligonucleotides can also be delivered using liposomes or by electroporation techniques which are well known in the art.

The antisense nucleic acids of the present invention can also be used to design antibiotic compounds comprising nucleic acids which function by intracellular triple helix formation. Triple helix oligonucleotides are used to inhibit transcription from a genome. The sequences identified as required for proliferation in the present invention, or portions thereof, can be used as templates to inhibit microorganism gene expression in individuals infected with such organisms. Traditionally, homopurine sequences were considered the most useful for triple helix strategies. However, homopyrimidine sequences can also inhibit gene expression. Such homopyrimidine oligonucleotides bind to the major groove at homopurine:homopyrimidine sequences. Thus, both types of sequences based on the sequences of the present invention that are required for proliferation are contemplated for use as antibiotic compound templates.

The antisense oligonucleotides of this example employ the identified sequences of the present invention to induce bacterial cell death or at least bacterial stasis by inhibiting target nucleic acid transcription or translation. Antisense oligonucleotides containing from about 8 to 40 nucleotides of the sequences of the present invention have sufficient complementary to form a duplex with the target sequence under physiological conditions.

To kill bacterial cells or inhibit their growth, the antisense oligonucleotides are applied to the bacteria or to the target cells under conditions that facilitate their uptake. These conditions

include sufficient incubation times of cells and oligonucleotides so that the antisense oligonucleotides are taken up by the cells. In one embodiment, an incubation period of 7-10 days is sufficient to kill bacteria in a sample. An optimum concentration of antisense oligonucleotides is selected for use.

The concentration of antisense oligonucleotides to be used can vary depending on the type of bacteria sought to be controlled, the nature of the antisense oligonucleotide to be used, and the relative toxicity of the antisense oligonucleotide to the desired cells in the treated culture. Antisense oligonucleotides can be introduced to cell samples at a number of different concentrations preferably between $1x10^{-10}M$ to $1x10^{-4}M$. Once the minimum concentration that can adequately control gene expression is identified, the optimized dose is translated into a dosage suitable for use *in vivo*. For example, an inhibiting concentration in culture of $1x10^{-7}$ translates into a dose of approximately 0.6 mg/kg body weight. Levels of oligonucleotide approaching 100 mg/kg body weight or higher may be possible after testing the toxicity of the oligonucleotide in laboratory animals. It is additionally contemplated that cells from the subject are removed, treated with the antisense oligonucleotide, and reintroduced into the subject. This range is merely illustrative and one of skill in the art are able to determine the optimal concentration to be used in a given case.

After the bacterial cells have been killed or controlled in a desired culture, the desired cell population may be used for other purposes.

EXAMPLE 41

The following example demonstrates the ability of an *E. coli* antisense oligonucleotide to act as a bactericidal or bacteriostatic agent to treat a contaminated cell culture system. The application of the antisense oligonucleotides of the present invention are thought to inhibit the translation of bacterial gene products required for proliferation.

The antisense oligonucleotide of this example corresponds to a 30 base phophorothioate modified oligodeoxynucelotide complementary to a nucleic acid involved in proliferation, such as Molecule Number EcXA118 (SEQ ID NO: 1). A sense oligodeoxynucelotide complementary to the antisense sequence is synthesized and used as a control. The oligonucleotides are synthesized and purified according to the procedures of Matsukura, et al., Gene 72:343 (1988). The test oligonucleotides are dissolved in a small volume of autoclaved water and added to culture medium to make a 100 micromolar stock solution.

Human bone marrow cells are obtained from the peripheral blood of two patients and cultured according standard procedures well known in the art. The culture is contaminated with the K-12 strain of *E. coli* and incubated at 37°C overnight to establish bacterial infection.

The control and antisense oligonucleotide containing solutions are added to the contaminated cultures and monitored for bacterial growth. After a 10 hour incubation of culture and oligonucleotides, samples from the control and experimental cultures are drawn and analyzed for the translation of the target bacterial gene using standard microbiological techniques well

known in the art. The target *E. coli* gene is found to be translated in the control culture treated with the control oligonucleotide, however, translation of the target gene in the experimental culture treated with the antisense oligonucleotide of the present invention is not detected or reduced.

One way to determine if a gene is essential for proliferation in a host or virulence in a host is to construct a conditional allele of the gene an infectious organism. The host is then challenged with the organism under conditions in which the product of the gene is functional or non-functional or has reduced activity or where the gene product is absent or else present but at a reduced level. If the gene is essential for proliferation or virulence the infection of the host will be diminished or abolished under conditions in which the product of the gene is not functional or has reduced activity or where the gene product is absent or else present but at a reduced level.

Since expression of an antisense nucleic acid complementary to a gene required for proliferation also decreases the synthesis of the gene product, antisense nucleic acids may also be used to evaluate whether a gene is essential for the proliferation or virulence of an infectious organism in the host. In such methods, nucleic acids encoding an antisense molecule complementary to the desired target gene are introduced into the infectious organism. For exampe, plasmids comprising one of SEQ ID NOs: 1-93 or fragments thereof which inhibit proliferation, may be introduced into the infectious organism. In some embodiments, the antisense nucleic acid may be transcribed from the IPTG-inducible promoter in pLEX5BA or from other regulated promoters or vector systems.

E. coli is transformed with the nucleic acid encoding the antisense molecule by electroporation and grown in medium which selects for the presence of the vector from which the antisense nucleic acid is expressed. The essentiality of the target for each antisense nucleic acid is verified in microorganisms grown in culture using the techniques described herein.

The ability of antisense expression to block *E. coli* infection in an animal is tested using the rabbit model of bacterial meningitis. A spinal needle is surgically placed into the cisterna magna of New Zealand White rabbits. The rabbits are inoculated with 10^5 to 10^6 cells of a normally virulent *E. coli* strain expressing an antisense nucleic acid complementary to a gene required for proliferation. Repeated CSF sampling is undertaken to determine multiple parameters of injury and infection such as cytochemical abnormalities, intracranial pressure, cerebral edema, BBB permeability, cerebral perfusion pressure and recovery of viable *E. coli* cells. Control animals are given intravenous injections of saline, which will not induce expression of the antisense nucleic acid, while experimental animals are given IPTG in intravenous injections to induce expression of the antisense nucleic acid. Alternatively, expression of the antisense nucleic acid may be induced by intravenous infusion of IPTG at sub-toxic levels. If other promoters other than IPTG inducible promoters are used, the rabbits may be fed the inducer in their water.

The use of rabbits allows multiple CSF samples per animal (one rabbit can give up to 8 sequential samples without change in CSF pressure). Treated animals receive therapy from 2 hours

post receiving inoculation up to several days. A typical efficacy study consists of 3 control animals and 3 treated animals.

The control animals in which expression of the antisense nucliec acid is not induced are not protected against infection with *E. coli* and there is a logarithmic increase in viable bacteria. In experimental animals, *E. coli* cells recovered from the site of infection are viable until antisense expression is subsequently induced. However, if the antisense nucleic acid is directed against a gene required for proliferation, after treatment with the inducer for antisense expression the *E. coli* cells infecting these rabbits will not multiply and fewer viable cells will be recovered from the site of infection. The *E. coli* cells recovered from the rabbits treated with the inducer are recovered, if still present, and assayed as above to determine if the promoter and gene are still present and functional. Conversely, if the antisense nucleic acid is not complementary to a gene required for proliferation, treatment of the rabbits with inducer will have no effect on *E. coli* viability.

EXAMPLE 42

A subject suffering from an *E. coli* infection is treated with the antisense oligonucleotide preparation of Example 39. The antisense oligonucleotide is provided in a pharmaceutically acceptable carrier at a concentration effective to inhibit the transcription or translation of the target nucleic acid. The present subject is treated with a concentration of antisense oligonucleotide sufficient to achieve a blood concentration of about 100 micromolar. The patient receives daily injections of antisense oligonucleotide to maintain this concentration for a period of 1 week. At the end of the week a blood sample is drawn and analyzed for the presence or absence of the orgranism using standard techniques well known in the art. There is no detectable evidence of E. coli and the treatment is terminated.

EXAMPLE 43

Preparation and use of Triple Helix Probes

The sequences of microorganism genes required for proliferation of the present invention are scanned to identify 10-mer to 20-mer homopyrimidine or homopurine stretches that could be used in triple-helix based strategies for inhibiting gene expression. Following identification of candidate homopyrimidine or homopurine stretches, their efficiency in inhibiting gene expression is assessed by introducing varying amounts of oligonucleotides containing the candidate sequences into a population of bacterial cells that normally express the target gene. The oligonucleotides may be prepared on an oligonucleotide synthesizer or they may be purchased commercially from a company specializing in custom oligonucleotide synthesis, such as GENSET, Paris, France.

The oligonucleotides can be introduced into the cells using a variety of methods known to those skilled in the art, including but not limited to calcium phosphate precipitation, DEAE-Dextran, electroporation, liposome-mediated transfection or native uptake.

Treated cells are monitored for a reduction in proliferation using techniques such as monitoring growth levels as compared to untreated cells using optical density measurements. The

oligonucleotides that are effective in inhibiting gene expression in cultured cells can then be introduced *in vivo* using the techniques well known in that art at a dosage level shown to be effective.

In some embodiments, the natural (beta) anomers of the oligonucleotide units can be replaced with alpha anomers to render the oligonucleotide more resistant to nucleases. Further, an intercalating agent such as ethidium bromide, or the like, can be attached to the 3′ end of the alpha oligonucleotide to stabilize the triple helix. For information on the generation of oligonucleotides suitable for triple helix formation see Griffin et al. (Science 245:967-971 (1989)).

EXAMPLE 44

Identification of Bacterial Strains from Isolated Specimens by PCR

Classical bacteriological methods for the detection of various bacterial species are time consuming and costly. These methods include growing the bacteria isolated from a subject in specialized media, cultivation on selective agar media, followed by a set of confirmation assays that can take from 8 to 10 days or longer to complete. Use of the identified sequences of the present invention provides a method to dramatically reduce the time necessary to detect and identify specific bacterial species present in a sample.

In one exemplary method, bacteria are grown in enriched media and DNA samples are isolated from specimens of, for example, blood, urine, stool, saliva or central nervous system fluid by conventional methods. A panel of PCR primers based on identified sequences unique to various species of microorganisms are then utilized in accordance with Example 12 to amplify DNA of approximately 100-200 nucleotides in length from the specimen. A separate PCR reaction is set up for each pair of PCR primers and after the PCR reaction is complete, the reaction mixtures are assayed for the presence of PCR product. The presence or absence of bacteria from the species to which the PCR primer pairs belong is determined by the presence or absence of a PCR product in the various test PCR reaction tubes.

Although the PCR reaction is used to assay the isolated sample for the presence of various bacterial species, other assays such as the Southern blot hybridization are also contemplated.

WHAT IS CLAIMED IS:

1. A purified or isolated nucleic acid sequence consisting essentially of one the sequence of nucleotides of SEQ ID NOs: 1-93, wherein expression of said nucleic acid in a microorganism is capable of inhibiting proliferation of a microorganism.

2. The nucleic acid sequence of Claim 1, wherein said nucleic acid sequence is complementary to at least a portion of the nucleotide sequence of the coding strand of a gene whose expression is required for proliferation of a microorganism.

3. The nucleic acid of Claim 1, wherein said nucleic acid sequence has a nucleotide seuqence complementary to at least a portion of the nucleotide sequence of an RNA required for proliferation of a microorganism.

4. The nucleic acid of Claim 3, wherein the nucleotide sequence of said RNA encodes more than one gene product.

5. A purified or isolated nucleic acid comprising a fragment of one of the nucleotide seuqence of SEQ ID NOs.: 1-93, said fragment selected from the group consisting of fragments comprising at least 10, at least 20, at least 25, at least 30, at least 50 and more than 50 consecutive nucleotides of one of the nucleotide sequences of SEQ ID NOs: 1-93.

6. A vector comprising a promoter operably linked to the nucleic acid sequence of Claims 1,2,3,4, or 5.

7. The vector of Claim 6, wherein said promoter is active in a microorganism selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling within the genera of any of the above species.

8. A host cell containing the vector of Claim 6.

9. A purified or isolated nucleic acid consisting essentially of the coding sequence of one of SEQ ID NOs: 106-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286.

10. A fragment of the nucleic acid of Claim 8, said fragment comprising at least 10, at least 20, at least 25, at least 30, at least 50 or more than 50 consecutive nucleotides of one of SEQ ID NOs: 106-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286.

11. A vector comprising a promoter operably linked to the nucleic acid of Claim 9 or Claim 10.

12. A purified or isolated antisense nucleic acid comprising a nucleic acid sequence complementary to at least a portion of an intragenic sequence, intergenic sequence, sequences spanning at least a portion of two or more genes, 5' noncoding region, or 3' noncoding region within an operon comprising a proliferation-required gene whose activity or expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93.

13. A purified or isolated nucleic acid comprising a nucleic acid having at least 70% identity to a sequence selected from the group consisting of SEQ ID NOs.: 1-93, fragments comprising at least 25 consecutive nucleotides of SEQ ID NOs.: 1-93, the sequences complementary to SEQ ID NOs.: 1-93 and the sequences complementary to fragments comprising at least 25 consecutive nucleotides of SEQ ID NOs.: 1-93 as determined using BLASTN version 2.0 with the default parameters.

14. The nucleic acid of Claim 13, wherein said nucleic acid is from an organism selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling within the genera of any of the above species.

15. A vector comprising a promoter operably linked to a nucleic acid encoding a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93.

16. A host cell containing the vector of Claim 15.

17. The vector of Claim 15, wherein said polypeptide comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

18. A purified or isolated polypeptide comprising a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93, or a fragment selected from the group consisting of fragments comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of one of the said polypeptides.

19. The polypeptide of Claim 18, wherein said polypeptide comprises a polypeptide comprising one of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 or a fragment comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

20. A purified or isolated polypeptide comprising a polypeptide having at least 25% identity to a polypeptide whose expression is inhibited by a sequence selected from the group consisting of SEQ ID NOs.: 1-93, or at least 25% identity to a fragment comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide whose expression is inhibited by a nucleic acid selected from the group consisting of SEQ ID NOs.: 1-93 as determined using FASTA version 3.0t78 with the default parameters.

21. The polypeptide of Claim 20, wherein said polypeptide has at least 25% identity to a polypeptide comprising one of SEQ ID NOs: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 or at least 25% identity to a fragment comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide comprising one of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 as determined using FASTA version 3.0t78 with the default parameters.

22. An antibody capable of specifically binding the polypeptide of one of Claims 18-21.

23. A method of producing a polypeptide, comprising introducing a vector comprising a promoter operably linked to a nucleic acid encoding a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93 into a cell and expressing said polypeptide.

24. The method of Claim 23, further comprising the step of isolating said polypeptide.

25. The method of Claim 23, wherein said polypeptide comprises a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

26. A method of inhibiting proliferation of a microorganism comprising inhibiting the activity or reducing the amount of a gene product whose expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 or inhibiting the activity or reducing the amount of a nucleic acid encoding said gene product.

27. The method of Claim 26, wherein said gene product comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

28. A method for identifying a compound which influences the activity of a gene product required for proliferation, said gene product comprising a gene product whose expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising:

contacting said gene product with a candidate compound; and

determining whether said compound influences the activity of said gene product.

29. The method of Claim 28, wherein said gene product is a polypeptide and said activity is an enzymatic activity.

30. The method of Claim 28, wherein said gene product is a polypeptide and said activity is a carbon compound catabolism activity.

31. The method of Claim 28, wherein said gene product is a polypeptide and said activity is a biosynthetic activity.

32. The method of Claim 28, wherein said gene product is a polypeptide and said activity is a transporter activity.

33. The method of Claim 28, wherein said gene product is a polypeptide and said activity is a transcriptional activity.

34. The method of Claim 28, wherein said gene product is a polypeptide and said activity is a DNA replication activity.

35. The method of Claim 28, wherein said gene product is a polypeptide and said activity is a cell division activity.

36. A compound identified using the method of Claim 28.

37. The method of Claim 28, wherein said gene product is a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

38. A method for identifying a compound or nucleic acid having the ability to reduce the activity or level of a gene product required for proliferation, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising:

(a) providing a target that is a gene or RNA, wherein said target comprises a nucleic acid encoding said gene product;

(b) contacting said target with a candidate compound or nucleic acid; and

(c) measuring an activity of said target.

39. The method of Claim 38, wherein said target is a messenger RNA molecule and said activity is translation of said messenger RNA.

40. The method of Claim 38, wherein said target is a messenger RNA molecule and said activity is transcription of a gene encoding said messenger RNA.

41. The method of Claim 38, wherein said target is a gene and said activity is transcription of said gene.

42. The method of Claim 38, wherein said target is a nontranslated RNA and said activity is processing or folding of said nontranslated RNA or assembly of said nontranslated RNA into a protein/RNA complex.

43. The method of Claim 38, wherein said target gene or RNA encodes a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

44. A compound or nucleic acid identified using the method of Claim 38.

45. A method for identifying a compound which reduces the activity or level of a gene product required for proliferation of a microorganism, wherein the activity or expression of said gene product is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising the steps of:

(a) expressing a sub-lethal level of an antisense nucleic acid complementary to a nucleic acid encoding said gene product in a cell to reduce the activity or amount of said gene product in said cell, thereby producing a sensitized cell;

(b) contacting said sensitized cell with a compound; and

(c) determining whether said compound inhibits the growth of said sensitized cell.

46. The method of Claim 45, wherein said determining step comprises determining whether said compound inhibits the growth of said sensitized cell to a greater extent than said compound inhibits the growth of a nonsensitized cell.

47. The method of Claim 45, wherein said cell is selected from the group consisting of bacterial cells, fungal cells, plant cells, and animal cells.

48. The method of Claim 45, wherein said cell is a Gram negative bacterium.

49. The method of Claim 45, wherein said cell is an *E. coli* cell.

50. The method of Claim 45, wherein said cell is from an organism selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling within the genera of any of the above species.

51. The method of Claim 45, wherein said antisense nucleic acid is transcribed from an inducible promoter.

52. The method of Claim 51, further comprising the step of contacting said cell with a concentration of inducer which induces said antisense nucleic acid to a sub-lethal level.

53. The method of Claim 45, wherein growth inhibition is measured by monitoring optical density of a culture growth solution.

54. The method of Claim 45, wherein said gene product is a polypeptide.

55. The method of Claim 54, wherein said polypeptide comprises a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

56. The method of Claim 45, wherein said gene product is an RNA.

57. A compound identified using the method of Claim 45.

58. A method for inhibiting cellular proliferation comprising introducing a compound with activity against a gene whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 or a compound with activity against the product of said gene into a population of cells expressing said gene.

59. The method of Claim 58, wherein said compound is an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, or a proliferation-inhibiting portion thereof.

60. The method of Claim 59, wherein said proliferation inhibiting portion of one of SEQ ID NOs.: 1-93 is a fragment comprising at least 10, at least 20, at least 25, at least 30, at least 50 or more than 51 consecutive nucleotides of one of SEQ ID NOs.: 1-93.

61. The method of Claim 58, wherein said population is a population selected from the group consisting of bacterial cells, fungal cells, plant cells, and animal cells.

62. The method of Claim 58, wherein said population is a population of Gram negative bacteria.

63. The method of Claim 58, wherein said population is a population of *E. coli* cells.

64. The method of Claim 58, wherein said population is a population selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae,Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling within the genera of any of the above species.

65. The method of Claim 58, wherein said gene encodes a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

66. A preparation comprising an effective concentration of an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, or a proliferation-inhibiting portion thereof in a pharmaceutically acceptable carrier.

67. The preparation of Claim 66, wherein said proliferation-inhibiting portion of one of SEQ ID NOs.: 1-93 comprises at least 10, at least 20, at least 25, at least 30, at least 50 or more than 50 consecutive nucleotides of one of SEQ ID NOs.: 1-93.

68. A method for inhibiting the activity or expression of a gene in an operon required for proliferation wherein the activity or expression of at least one gene in said operon is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising contacting a cell in a cell population with an antisense nucleic acid comprising at least a proliferation-inhibiting portion of said operon.

69. The method of Claim 68, wherein said antisense nucleic acid comprises a sequence selected from the group consisting of SEQ ID NOs.: 1-93 or a proliferation inhibiting portion thereof.

70. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a plasmid which expresses said antisense nucleic acid into said cell population.

71. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a phage which expresses said antisense nucleic acid into said cell population.

72. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by expressing said antisense nucleic acid from the chromosome of cells in said cell population.

73. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a promoter adjacent to a chromosomal copy of said antisense nucleic acid such that said promoter directs the synthesis of said antisense nucleic acid.

74. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a retron which expresses said antisense nucleic acid into said cell population.

75. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a ribozyme into said cell-population, wherein a binding portion of said ribozyme is complementary to said antisense oligonucleotide.

76. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a liposome comprising said antisense oligonucleotide into said cell.

77. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by electroporation of said antisense nucleic acid.

78. The method of Claim 68, wherein said antisense nucleic acid is a fragment comprising at least 10, at least 20, at least 25, at least 30, at least 50 or more than 50 consecutive nucleotides of one of SEQ ID NOs.: 1-93.

79. The method of Claim 68 wherein said antisense nucleic acid is an oligonucleotide.

80. A method for identifying a gene which is required for proliferation of a microorganism comprising:

(a) contacting a microorganism other than *E. coli* with a nucleic acid selected from the group consisting of SEQ ID NOs.: 1-93;

(b) determining whether said nucleic acid inhibits proliferation of said microorganism; and

(c) identifying the gene in said microorganism which is inhibited by said nucleic acid.

81. The method of Claim 80, wherein said microorganism is a Gram negative bacterium.

82. The method of Claim 80 wherein said microorganism is selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus,*

Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis and any species falling within the genera of any of the above species.

83. The method of Claim 80, further comprising introducing said nucleic acid into a vector functional in said microorganism prior to introducing said inhibitory nucleic acid into said microorganism.

84. A method for identifying a compound having the ability to inhibit proliferation of a microorganism comprising:

(a) identifying in a first microorganism a homolog of a gene or gene product present in a second microorganism which is different than said first microorganism, wherein the activity or level of said gene or gene product is inhibited by a nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs. 1-93 ;

(b) identifying an inhibitory nucleic acid sequence which inhibits the activity of said homolog in said first microorganism;

(c) contacting said first microorganism with a sub-lethal level of said inhibitory nucleic acid, thus sensitizing said first microorganism;

(d) contacting the sensitized microorganism of step (c) with a compound; and

(e) determining whether said compound inhibits proliferation of said sensitized microorganism.

85. The method of Claim 84, wherein said determining step comprises determining whether said compound inhibits proliferation of said sensitized microorganism to a greater extent than said compound inhibits proliferation of a nonsensitized microorganism.

86. The method of Claim 84 wherein step (a) comprises identifying a homologous nucleic acid to a gene or gene product whose activity or level is inhibited by a nucleic acid selected from the group consisting of SEQ ID NOs. 1-93 or a nucleic acid encoding a homologous polypeptide to a polypeptide whose activity or level is inhibited by a nucleic acid selected from the group consisting of SEQ ID NOs. 1-93 by using an algorithm selected from the group consisting of BLASTN version 2.0 with the default parameters and FASTA version 3.0t78 algorithm with the default parameters to identify said homologous nucleic acid or said nucleic acid encoding a homologous polypeptide in a database.

87. The method of Claim 84 wherein said step (a) comprises identifying a homologous nucleic acid or a nucleic acid encoding a homologous polypeptide by identifying nucleic acids which hybridize to said first gene.

88. The method of Claim 84 wherein the step (a) comprises expressing a nucleic acid selected from the group consisting of SEQ ID NOs. 1-93 in said microorganism.

89. The method of Claim 84, wherein said inhibitory nucleic acid is an antisense nucleic acid.

90. The method of Claim 84, wherein said inhibitory nucleic acid comprises an antisense nucleic acid to a portion of said homolog.

91. The method of Claim 84, wherein said inhibitory nucleic acid comprises an antisense nucleic acid to a portion of the operon encoding said homolog.

92. The method of Claim 84, wherein the step of contacting the first microorganism with a sub-lethal level of said inhibitory nucleic acid comprises directly contacting said microorganism with said inhibitory nucleic acid.

93. The method of Claim 84, wherein the step of contacting the first microorganism with a sub-lethal level of said inhibitory nucleic acid comprises expressing an antisense nucleic acid to said homolog in said microorganism.

94. The method of Claim 84, wherein said gene product comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

95. A compound identified using the method of Claim 84.

96. A method of identifying a compound having the ability to inhibit proliferation comprising:

(a) contacting a microorganism other than *E. coli* with a sub-lethal level of a nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs. 1-93 or a portion thereof which inhibits the proliferation of *E. coli*, thus sensitizing said microorganism;

(b) contacting the sensitized microorganism of step (a) with a compound; and

(c) determining whether said compound inhibits proliferation of said sensitized microorganism.

97. The method of Claim 96, wherein said determining step comprises determining whether said compound inhibits proliferation of said sensitized microorganism to a greater extent than said compound inhibits proliferation of a nonsensitized microorganism.

98. A compound identified using the method of Claim 96.

99. A method for identifying a compound having activity against a biological pathway required for proliferation comprising:

(a) sensitizing a cell by expressing a sub-lethal level of an antisense nucleic acid complementary to a nucleic acid encoding a gene product required for proliferation, wherein the activity or expression of said gene product is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, in said cell to reduce the activity or amount of said gene product;

(b) contacting the sensitized cell with a compound; and

(c) determining whether said compound inhibits the growth of said sensitized cell.

100. The method of Claim 99, wherein said determining step comprises determining whether said compound inhibits the growth of said sensitized cell to a greater extent than said compound inhibits the growth of a nonsensitized cell.

101. The method of Claim 99, wherein said cell is selected from the group consisting of bacterial cells, fungal cells, plant cells, and animal cells.

102. The method of Claim 99, wherein said cell is a Gram negative bacterium.

103. The method of Claim 99, wherein said Gram negative bacterium is *E. coli.*

104. The method of Claim 99, wherein said cell is selected from the group consisting of *Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis* and any species falling within the genera of any of the above species.

105. The method of Claim 99, wherein said antisense nucleic acid is transcribed from an inducible promoter.

106. The method of Claim 99, further comprising contacting the cell with an agent which induces expression of said antisense nucleic acid from said inducible promoter, wherein said antisense nucleic acid is expressed at a sub-lethal level.

107. The method of Claim 99, wherein inhibition of proliferation is measured by monitoring the optical density of a liquid culture.

108. The method of Claim 99, wherein said gene product comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479

109. A compound identified using the method of Claim 99.

110. A method for identifying a compound having the ability to inhibit cellular proliferation comprising:

(a) contacting a cell with an agent which reduces the activity or level of a gene product required for proliferation of said cell, wherein said gene product is a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93;

(b) contacting said cell with a compound; and

(c) determining whether said compound reduces proliferation of said contacted cell.

111. The method of Claim 110, wherein said determining step comprises determining whether said compound reduces proliferation of said contacted cell to a greater extent than said compound reduces proliferation of cells which have not been contacted with said agent.

112. The method of Claim 110, wherein said agent which reduces the activity or level of a gene product required for proliferation of said cell comprises an antisense nucleic acid to a gene or operon required for proliferation.

113. The method of Claim 110, wherein said agent which reduces the activity or level of a gene product required for proliferation of said cell comprises a compound known to inhibit growth or proliferation of a microorganism.

114. The method of Claim 110, wherein said cell contains a mutation which reduces the activity or level of said gene product required for proliferation of said cell.

115. The method of Claim 114, wherein said mutation is a temperature sensitive mutation.

116. The method of Claim 110, wherein said gene product comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479

117. A compound identified using the method of Claim 110.

118. A method for identifying the biological pathway in which a proliferation-required gene or its gene product lies, wherein said gene or gene product comprises a gene or gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising:

(a) expressing a sub-lethal level of an antisense nucleic acid which inhibits the activity of said proliferation-required gene or gene product in a cell;

(b) contacting said cell with a compound known to inhibit growth or proliferation of a microorganism, wherein the biological pathway on which said compound acts is known; and

(c) determining whether said cell is sensitive to said compound.

119. The method of Claim 118, wherein said determining step comprises determining whether said cell has a substantially greater sensitivity to said compound than a cell which does not express said sub-lethal level of said antisense nucleic acid and wherein said gene or gene product lies in the same pathway on which said compound acts if said cell expressing said sub-lethal level of said antisense nucleic acid has a substantially greater sensitivity to said compound than said cell which does not express said sub-lethal level of said antisense nucleic acid.

120. The method of Claim 118, wherein said gene product comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479

121. A method for determining the biological pathway on which a test compound acts comprising:

(a) expressing a sub-lethal level of an antisense nucleic acid complementary to a proliferation-required nucleic acid in a cell, wherein the activity or expression of said proliferation-required nucleic acid is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 and wherein the biological pathway in which said proliferation-required nucleic acid or a protein encoded by said proliferation-required polypeptide lies is known,

(b) contacting said cell with said test compound; and

(c) determining whether said cell is sensitive to said test compound.

122. The method of Claim 121, wherein said determining step comprises determining whether said cell has a substantially greater sensitivity to said test compound than a cell which does not express said sub-lethal level of said antisense nucleic acid.

123. The method of Claim 121, further comprising:

(d) expressing a sub-lethal level of a second antisense nucleic acid complementary to a second proliferation-required nucleic acid in a second cell, wherein said second proliferation-required nucleic acid is in a different biological pathway than said proliferation-required nucleic acid in step (a); and

(e) determining whether said second cell does not have a substantially greater sensitivity to said test compound than a cell which does not express said sub-lethal level of said second antisense nucleic acid, wherein said test compound is specific for the biological pathway against which the antisense nucleic acid of step (a) acts if said second cell does not have substantially greater sensitivity to said test compound.

124. A purified or isolated nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93.

125. A compound which interacts with a gene or gene product whose activity or expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93 to inhibit proliferation.

126. A compound which interacts with a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93 to inhibit proliferation.

127. A method for manufacturing an antibiotic comprising the steps of:

screening one or more candidate compounds to identify a compound that reduces the activity or level of a gene product required for proliferation, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93; and

manufacturing the compound so identified.

128. The method of Claim 127, wherein said screening step comprises performing any one of the methods of Claims 28, 38, 45, 96, 99 and 110.

129. A method for inhibiting proliferation of a microorganism in a subject comprising administering a compound that reduces the activity or level of a gene product required for proliferation of said microorganism, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 to said subject.

130. The method of Claim 129 wherein said subject is selected from the group consisting of vertebrates, mammals, avians, and human beings.

131. The method of Claim 129, wherein said gene product comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 299-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

1/4

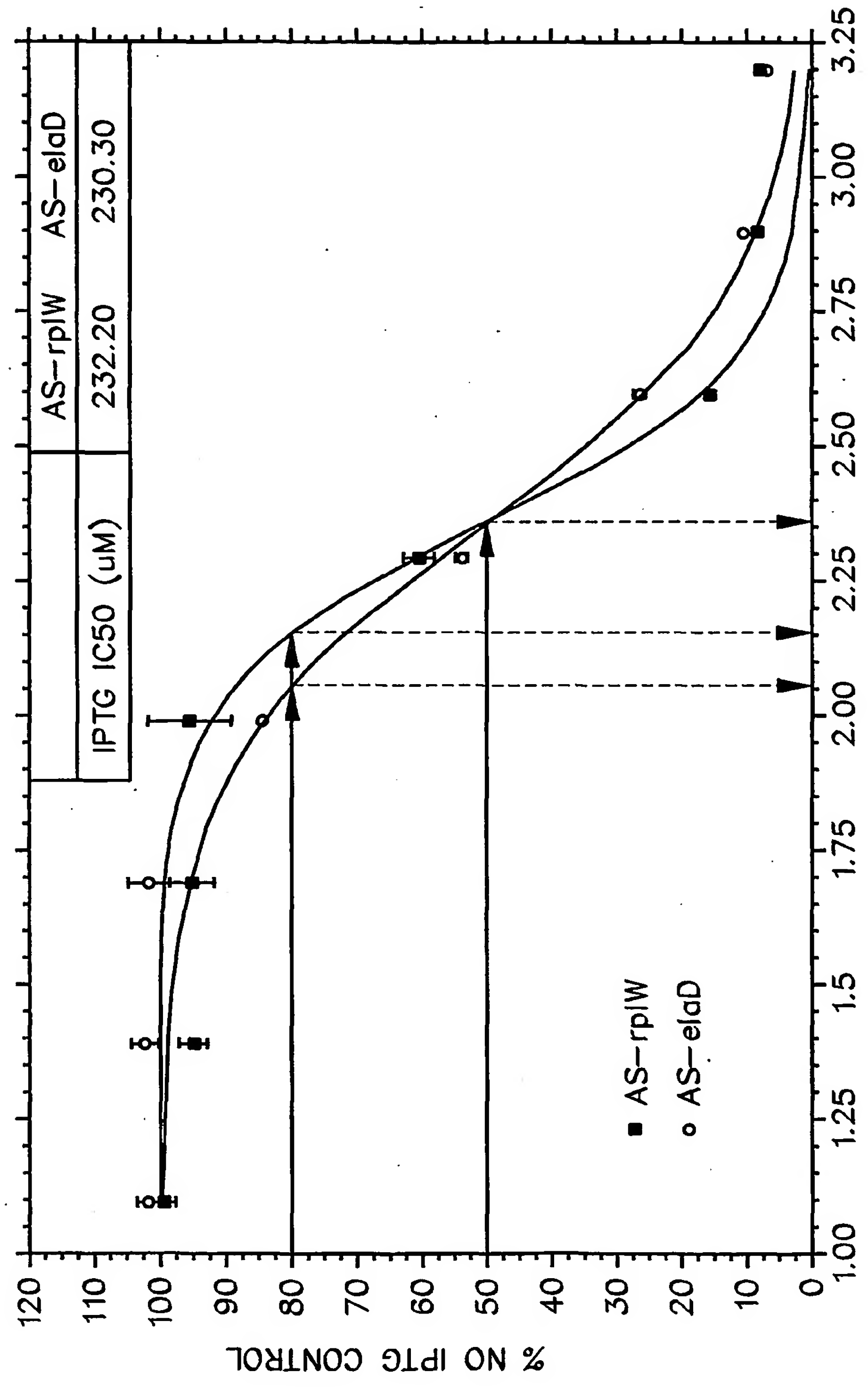

IPTG IC50 (uM)
AS—rplW 232.20
AS—elaD 230.30
■ AS—rplW
○ AS—elaD
% NO IPTG CONTROL
LOG(uM IPTG)
0
10
20
30
40
50
60
70
80
90
100
110
120
1.00
1.25
1.5
1.75
2.00
2.25
2.50
2.75
3.00
3.25

FIG. 1

2/4

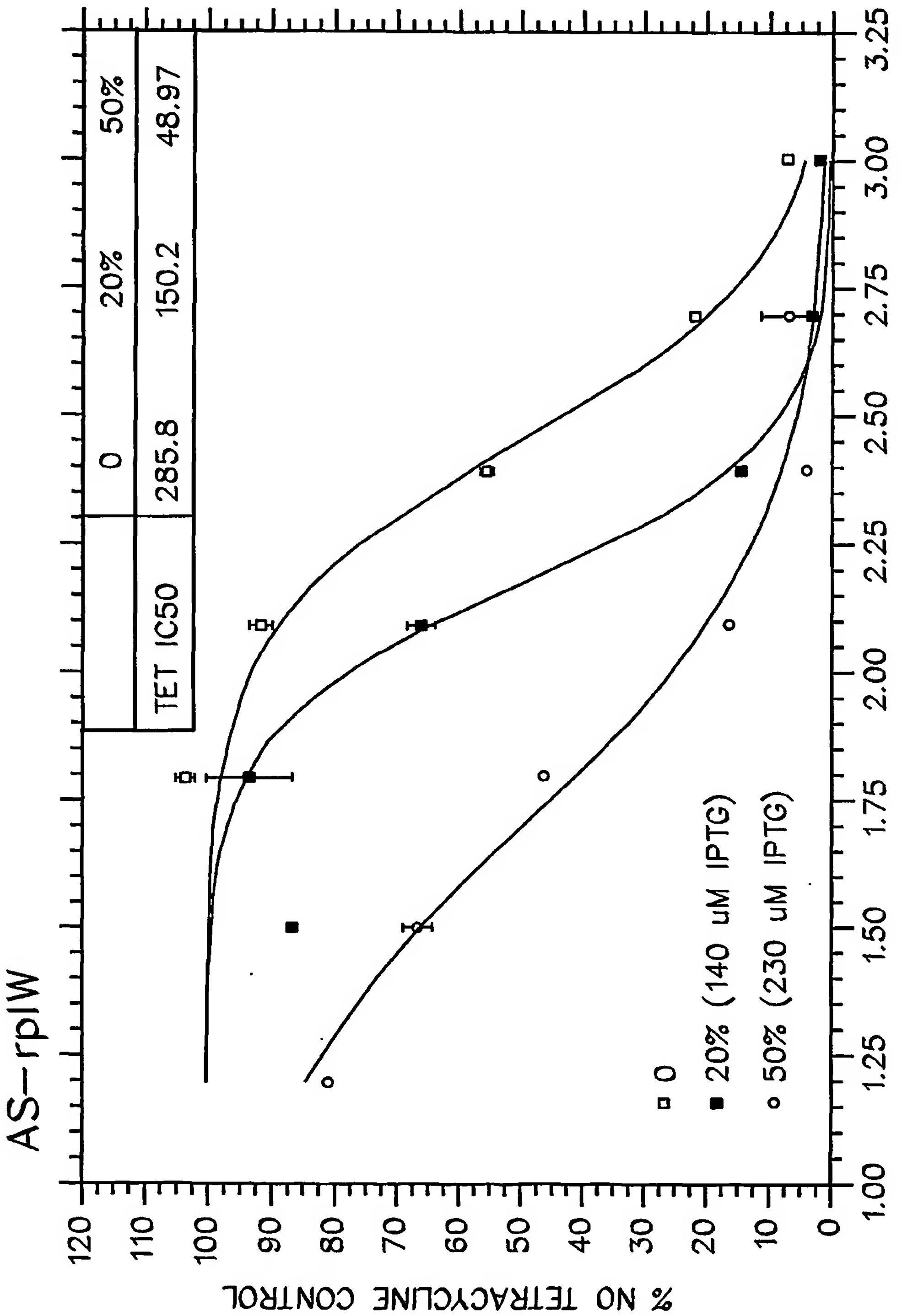
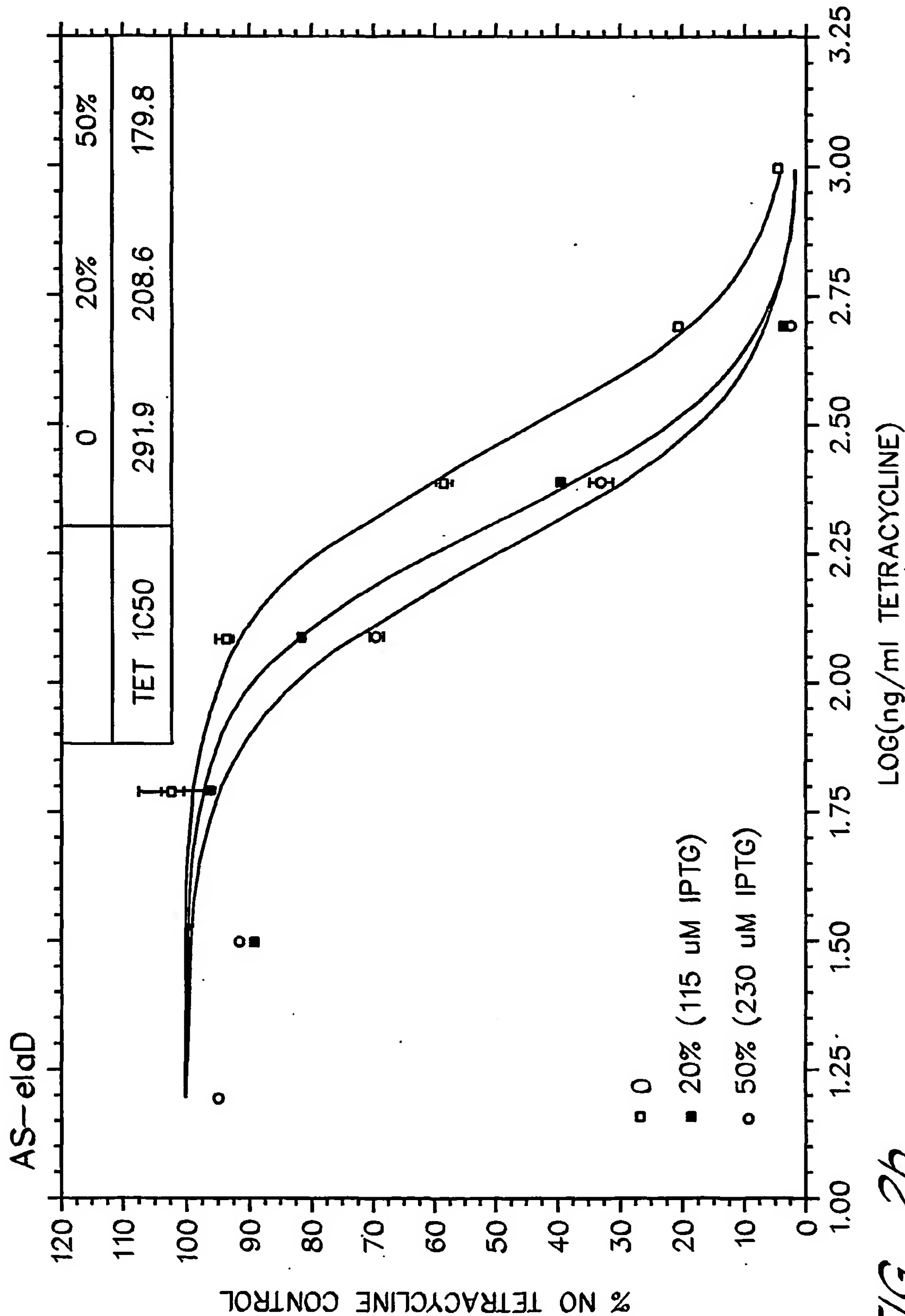

AS—rplW
TET IC50
0
285.8
20%
150.2
50%
48.97
□ 0
■ 20% (140 uM IPTG)
○ 50% (230 uM IPTG)
% NO TETRACYCLINE CONTROL
LOG(ng/ml TETRACYCLINE)
120
110
100
90
80
70
60
50
40
30
20
10
0
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
3.25

FIG. 2a

3/4



AS-elaD
TET 1C50
0
291.9
20%
208.6
50%
179.8
□ 0
■ 20% (115 uM IPTG)
○ 50% (230 uM IPTG)
% NO TETRACYCLINE CONTROL
120
110
100
90
80
70
60
50
40
30
20
10
0
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
3.25
LOG(ng/ml TETRACYCLINE)

FIG. 2b

4/4

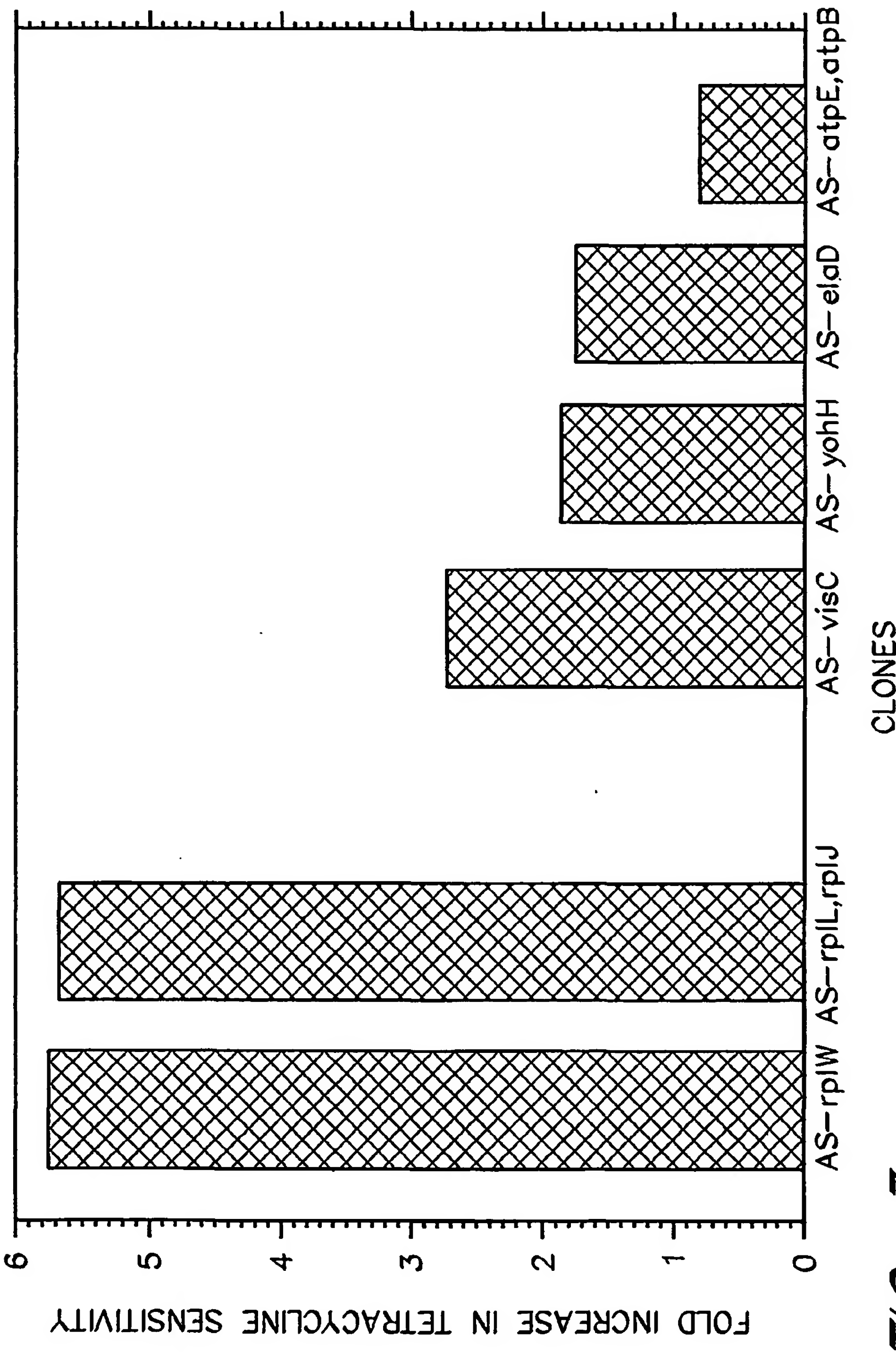

FOLD INCREASE IN TETRACYCLINE SENSITIVITY
0
1
2
3
4
5
6
AS–rplW
AS–rplL,rplJ
AS–visC
AS–yohH
AS–elaD
AS–atpE,atpB
CLONES

FIG. 3

SEQUENCE LISTING

<110> ELITRA PHARMACEUTICALS, INC.
FORSYTH, R. Allyn
OHLSEN, Kari L.
ZYSKIND, Judith W.

<120> GENES IDENTIFIED AS REQUIRED FOR
PROLIFERATION OF E. COLI

<130> ELITRA.009VPC

<150> US 60/173,005
<151> 1999-12-23

<160> 481

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 142
<212> DNA
<213> Escherichia coli

<400> 1
cgctccttag agtacttctg gtgccagaga gataattttc atgaacttct cactacgaag 60
ctcacgagtt accggcccaa aaatacgcgt accgataggc tgctcgctgt tgttgttcag 120
aagaacacaa gcattaccat cg 142

<210> 2
<211> 154
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)...(154)
<223> n = A,T,C or G

<400> 2
gctccttaga gtacttntgg tgccggagag ataattttca tgaacttctc actacnaagc 60
tcacgagtta ccggcccaaa aatacgcgta ccgataggct gctcgctgtt gttgttcacg 120
aagaacacaa gcattaccat cgaagcgaat gaca 154

<210> 3
<211> 140
<212> DNA
<213> Escherichia coli

<400> 3
tacttctggt gccagagaga taattttcat gaacttctca ctacgaagct cacgagttac 60
cggcccaaaa atacgcgtac cgataggctg ctcgctgttg ttgttcagaa gaacacaagc 120
attaccatcg aagcgaatga 140

<210> 4
<211> 177
<212> DNA
<213> Escherichia coli

<400> 4
cgctccttag agtacttctg gtgccagaga gataattttc atgaacttct cactacgaag 60

ctcacgagtt accggcccaa aaatacgcgt accgataggc tgctcgctgt tgttgttcag 120
aagaacacaa gcattaccat cgaagcgaat gacagaaccg tccgggcgac gaacacc 177

<210> 5
<211> 255
<212> DNA
<213> Escherichia coli

<400> 5
ctctttttct tcgagatgga cgcagtgagc aaagacacag tttttaccgg tcaggccgta 60
ctgatggtaa acatccagat aaccatcatg gtcaggataa agcgatttca cccaggcaat 120
ttcatctttg ttttcacaga gatgggtatg tacccacgta tccggatatt cttctttcag 180
gcgttgcgcc atcgccatct gttcaggaga tgaggtcggg gcgaagcgtg gcgtaatcgc 240
atatagcaga cgacc 255

<210> 6
<211> 395
<212> DNA
<213> Escherichia coli

<400> 6
tccgcgatcg cggataaaaa ggaggattta aataaggcgt tgatgcgccg tttgaagtcg 60
gcaatattaa gccgcatgcc atctcgacat gcggcttata cggtttacca gctatatttc 120
actcccagca tgccctgagt atcgctatag cctttatcac ctagttgcac accgacattc 180
ccccacaggc taaggttgtt atttactttc gcctcaaccc cggtacgtac ttcaccgaga 240
ttacgcgcac cttcacggcc tacggtttga ccattcatct tcacggcgta gactttgcta 300
ttgttgatcc agttcgcttc aatgtaaggc tggaactcac gctgtttacc atcgtcacgc 360
tggtgatggc tgttcaggta ggttttcacc ccaag 395

<210> 7
<211> 280
<212> DNA
<213> Escherichia coli

<400> 7
tgatcgatca ggtaatttgc cagcggtaga taattagtga tgatttgcac tggcttgcca 60
cacatttccc gcccaagcag aaacgcggtg gagccgcagt tgatgactac gctttcgccc 120
ggattaacca gctgcgaggc cgctttagcg atacgtactt tttcatcgtg attctgcgcc 180
tgatgcagat tcatcggcgt ccagcgcggg cgctgttggg taatagcttc tgcgccattg 240
cgcacttttt tcagtttgcc gctttcgtca agtttattga 280

<210> 8
<211> 363
<212> DNA
<213> Escherichia coli

<400> 8
agtaaaggat atacaaattt aaaggaagtt gcggcgaagt caaaacttac agatacaacg 60
gtttcaattc cgataacagc cagttactac gtctacgata caaacaaagt taaatctggc 120
gcactggagg caaccgcatt aatcaacgtg aaatacgact aataacaagc aagacgagca 180
agtggctaat aataaaaaat aacaaggtgt ctgcatttcc ccctcatgat gaggggctt 240
tttttagcga tgataaaaaa tctcaccgtc gtaggcttta atgattttac cgtcggtgtc 300
gctgatcaac acgtacgcgc cacccatata ggtccagtgc gtcccggcat caggcgcggg 360
cag 363

<210> 9
<211> 295
<212> DNA
<213> Escherichia coli

<400> 9

ttacagatac aacggtttca attccgataa cagccagtta ctacgtctac gatacaaaca 60
aagttaaatc tggcgcactg gaggcaaccg cattaatcaa cgtgaaatac gactaataac 120
aagcaagacg agcaagtggc taataataaa aaataacaag gtgtctgcat ttccccctca 180
tgatgagggg gctttttta gcgatgataa aaaatctcac cgtcgtaggc tttaatgatt 240
ttaccgtcgg tgtcgctgat caacacgtac gcgccaccca tataggtcca gtgcg 295

<210> 10
<211> 161
<212> DNA
<213> Escherichia coli

<400> 10
acgactaata acaagcaaga cgagcaagtg gctaataata aaaaataaca aggtgtctgc 60
atttccccct catgatgagg gggctttttt tagcgatgat aaaaaatctc accgtcgtag 120
gctttaatga ttttaccgtc ggtgtcgctg atcaacacgt a 161

<210> 11
<211> 290
<212> DNA
<213> Escherichia coli

<400> 11
tgcggcgaag tcaaaactta cagatacaac ggtttcaatt ccgataacag ccagttacta 60
cgtctacgat acaaacaaag ttaaatctgg cgcactggag gcaaccgcat taatcaacgt 120
gaaatacgac taataacaag caagacgagc aagtggctaa taataaaaaa taacaaggtg 180
tctgcatttc cccctcatga tgagggggct ttttttagcg atgataaaaa atctcaccgt 240
cgtaggcttt aatgatttta ccgtcggtgt cgctgatcaa cacgtacgcg 290

<210> 12
<211> 137
<212> DNA
<213> Escherichia coli

<400> 12
ctggttttag cttttgcttc gtaaacacgt aataaaacgt cctcacacaa tatgaggacg 60
ccgaatttta gggcgatgcc gaaaaggtgt caagaaatat acaacgatcc cgccatcacc 120
aggccatctg gctgggg 137

<210> 13
<211> 275
<212> DNA
<213> Escherichia coli

<400> 13
cgtcatgtac gtgcagtttg gtcgtacgct tgatgaattt accgtagatc gggtgtttca 60
caaaacgttc gatagcaaca acaatggatt tctccatttt gtcgctaaca acgcgacctt 120
gcagagtacg gattttatcg gtcattacgc acccgccttc tcgttcagta aagtcttaac 180
gcgtgcgaca tcgcgacgca cttgcttcaa caggtgagac tgttgcagct ggccacttgc 240
agcctgcata cgcaggttga actgctcacg cagca 275

<210> 14
<211> 239
<212> DNA
<213> Escherichia coli

<400> 14
tgtcggtaaa caatgacaat ttgagacaga gtgaaaggtc agatttgcgg agtaatgcac 60
ataatggtta tttaaataaa ccacatgaat cattaatggg ttattcatta tttttgtgat 120
ttactcacaa tgtatcaggg aaatataact taccgggaga tgtaatatgt ttaatttttc 180
aaatcgaaat ttaaaatatt gtgccggagg catctctggc acattgggca attacggca 239

<210> 15
<211> 114
<212> DNA
<213> Escherichia coli

<400> 15
cgatgttctg aatctgctcg cgcatctgct caatcaacac tttcagctcg atggcggagt 60
ttgtcacttc ggcattgata gacttcgacg caagagtgtt cgactcgcgg ttga 114

<210> 16
<211> 290
<212> DNA
<213> Escherichia coli

<400> 16
acgggaagta attgtaactt agtttgtaaa ggtcatatgc ttgttgatat ttctttttgg 60
gctgacatac tgatgcaaaa tcttttagat actcatgatt ttcgaacgca taaatgcaca 120
agaatttata ataaatttct gcgtccgtta tattaccctt ttcatggtag tcgaatgcat 180
atgcatatat tctatcaagg atatcaggag ttaaaccatg tagtaatgcc agattttctc 240
ctttcgataa tgcgtgtttt agttgatttg cccattcatc actattattg 290

<210> 17
<211> 444
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)...(444)
<223> n = A,T,C or G

<400> 17
acgggaagta attgtaactt agtttgtaaa ggtcatatgc ttgttgatat ttctttttgg 60
gctgacatac tgatgcaaaa tcttttagat actcatgatt ttcgaacgca taaatgcaca 120
agaatttata ataaatttct gcgtccgtta tattaccctt ttcatggtag tcgaatgcat 180
atgcatatat tctatcaagg atatcaggag ttaaaccatg tagtaatgcc agattttctc 240
ctttcgataa tgcgtgtttt agttgatttg cccattcatc actattantg aatatttcaa 300
ttgtttctgt gctcataata ttcctctgng atgtgnctcg ngatgccnaa cnattatttg 360
antgactcat ngnnattntt atatnnatcc nntnntaant natnantntn aagccngcnt 420
tgcttatggn ntantntnat ttat 444

<210> 18
<211> 250
<212> DNA
<213> Escherichia coli

<400> 18
cgctggaaaa aacctaaaat attcatctta ttccccctac gagaacccta tttggctcgt 60
ttcaagccgt atttttattt tgctgcaaat tgtactgccg atgttctgta atcagattgt 120
tagatcatct gctacagagt gtgtgaaaat ttaattcgta tcgcaaatta aacgcgtgtc 180
ttttgtgagt tttgtcacca aatatcgtta ttatcactcc cttttactgg ctaaaccaga 240
aaacttattt 250

<210> 19
<211> 237
<212> DNA
<213> Escherichia coli

<400> 19
gtattacaac aagaatacag tgctaaaaaa agaatgtacc tgaagatagt cttcattttt 60
actccaatgt ttctttattg ttatggtttg cgctgttttt gtttgttatt cgtagtaaat 120

atccactaac acatggctca taagttgacc agatgtaaca ctcaatgcag gatcgatttt 180
ttgcattcgg gtgctaaatt tccaggatga accattcatg tcatccgaca ttactgg 237

<210> 20
<211> 237
<212> DNA
<213> Escherichia coli

<400> 20
catcatcaca gcggcagcca tgtacagcag atccatattc aggttttcca tgacagtctc 60
cagtttgttt cagttaaaac gtagtagtgt tggtaaatta atgttcttca gacgccatcg 120
acagatagac gatcgtcaga accatgaaga tgaaggcttg cagcgtaatg atcaggatgt 180
ggaaaatggc ccacggcaca ttcaggatcc actgtgacca ccacggcaac agaccag 237

<210> 21
<211> 139
<212> DNA
<213> Escherichia coli

<400> 21
cccgcgacgt tatccgcgtg taggcatcgt catgacgtaa acggcgatcg gcggtataac 60
ggttgatcat cgaatccatg ttgccagcag taacaccgaa aaacagattc ggtttaccca 120
gacgcataaa gtcgtcttt 139

<210> 22
<211> 211
<212> DNA
<213> Escherichia coli

<400> 22
ggagagcgaa gggctatccg gccagggtga aattatcgcc gcgaacgcac aatttgatat 60
cgacgagtaa agtactcaaa cggcgcgctc cacacatgca ccagtcgggt gaacgggaag 120
agcaggaaga tggtcattcc caacaccaga tgcaggcgga agacgaacgc tacgccgttg 180
agcatctctg acgatccacc acggaaagtg a 211

<210> 23
<211> 154
<212> DNA
<213> Escherichia coli

<400> 23
cacgacggat tttcgctgcc atgattcgct ccttagagta cttctggtgc cagagagata 60
attttcatga acttctcact acgaagctca cgagttaccc ggcccaaaaa tacgcgtacc 120
gataggctgc tcgctgttgt tgttcagaag aaca 154

<210> 24
<211> 160
<212> DNA
<213> Escherichia coli

<400> 24
acttcgtcat cacgacggat tttcgctgcc atgattcgct ccttagagta cttctggtgc 60
cagagagata attttcatga acttctcact acgaagctca cgagttaccg gcccaaaaat 120
acgcgtaccg ataggctgct cgctgttgtt gttcagaaga 160

<210> 25
<211> 244
<212> DNA
<213> Escherichia coli

<400> 25

aatgaccttg ccggaagaca ggacattctt aactttaccg cgtttacctt tatctttacc 60
ggttaacacg ataacttcgt catcacgacg gattttcgct gccatgattc gctccttaga 120
gtacttctgg tgccagagag ataattttca tgaacttctc actacgaagc tcacgagtta 180
ccggcccaaa aatacgcgta ccgataggct gctcgctgtt gttgttcaga agaacacaag 240
catt 244

<210> 26
<211> 108
<212> DNA
<213> Escherichia coli

<400> 26
tgttcgatct cgaaacttac acctttctca gccagcacaa tgcggacctg atggctatag 60
atgtcagtag gaccggaaaa cagcgtcatt acccgaacgt ttgttggc 108

<210> 27
<211> 369
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)...(369)
<223> n = A,T,C or G

<400> 27
tccagacgct ctttcttggc gtcgtagtca aagatacccc ctaagaacgt cggagcgttc 60
cgtgagtcct gaatgcgatt atttaccgga ttaatttcaa acatggtctg atttctttta 120
ttgagctagt caaaatgcgg tgataagagc gggattgtac ccaatccacg ctctttttta 180
tagagaagat gacgctaaat tggccagata ttgtcgatga taatttgcag gctgcggttc 240
cacgaaactc gttgatatcg agcttataag ccagttgcac ttngcgcacc cgttatccgg 300
ccagagggcg gtatcgacat taaaagcaat accatccagc agtggaccgc cgcccgaccg 360
gttcgacca 369

<210> 28
<211> 457
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)...(457)
<223> n = A,T,C or G

<400> 28
acccccctaag aacgtcggag cgttccgtga ggtcctgaat gcgattattt accggattaa 60
tttcaaacat ggtctgattt cttttattga gctagtcaaa atgcggtgat aagagcggga 120
ttgtacccaa tccacgctct tttttataga gaagatgacg ctaaattggc cagatattgt 180
cgatgataat ttgcaggctg cggttgccgc gaaactcgtt gatatcgagc ttataagcca 240
gttgcacttc gcgcacgccg ttatccggcc agagggcggt atcgacatta aaagcaatac 300
catccagcag tggaccgccg ccgaccggtt cgaccatnac cttcaaatga cgttcgccca 360
ccagccgctg ttgcancaga cggaaatgac cgtcaaacag cggctccggg aacatctgcc 420
cccacgggcc agcatcgcgc agcagctgcg ccacttc 457

<210> 29
<211> 729
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature

<222> (1)...(729)
<223> n = A,T,C or G

<400> 29
ctttttacgg acaaacaatc atgcaaccag aagttgaagt attaaccgat cataatgagc 60
taatttgttc gagctttatt gaacacattg ccaacacatt aaatttagga gtggtgtaca 120
actaaacctg cagccaagga tgtatagtga gcgaagccct atcaggcctt tttggtcagt 180
agataagatt gatcttcgtt gatagaattt acttacacca gctgttacat taagataatt 240
ttttggtggg agaatgataa gatcttacgt aacaatttga ttttaatggt gccgataata 300
ggagtcgaac ctacgacctt cgcattacga atgcgctgct ctaccaactg agctatatcg 360
gccctgaaag gacatgttca cgaacgtgaa tcacggtgga caaggttaaa actaaccggg 420
cgatgcgtca atggccttgt gaatcaaatg gctacttttg catcacccgg ttttatttac 480
gcacgaatgg tgtaatcacc aatgccgatc cacttgtaag tggtcaagtg cttncagccc 540
cattgggcca cgcgccgtgg agtttttgtg tgcttaccgg cactttcgca cccagaccaa 600
actggccgcc gtcggtaaaa cgcgtagagg cgttacgtaa acagcggacg aatccacttc 660
gtnacaaaac gcttggcgtt ggcatatcgc gggtcangat cgcaatcgga gtgttgtgtg 720
cccgggttn 729

<210> 30
<211> 305
<212> DNA
<213> Escherichia coli

<400> 30
agccattgtc cggagaattt tcgtctttct cgagggtatt atctgaaagt aactcttctg 60
agctagagtc ttctatgttt aaatccatag tataattaat ttgggggacg accagataag 120
cattttcttt ggtattttca taatcttcaa tgattatatc cgatgtaatc atatattttt 180
tctgatgaga attctttcat ctctgttaat ttaattctgt tcaagcaatt gatgattttc 240
gatgagcttt tatatatttc ttttgttaat tcattttcta atttatctgt aagagttgcc 300
cttgg 305

<210> 31
<211> 425
<212> DNA
<213> Escherichia coli

<400> 31
aggttactgg cgaggttatc gagaaacgcc agactgatac ccataaaaaa gagaataagg 60
taaaccatcc agcttaatag ctgattaata acttttaacg cagcttgttg gcgaagcgga 120
atgaggtaac ccacaatcag gggaaccaga atgattaaca gcccagaaaa catgaaaacc 180
cagtccttgc aaagatgaag tcgaaatgcg cgatgacaca ctactgaaag cggaaggacg 240
agtaaagttg caattaaaag gaaatgttat gcataaggag cagtagagta ttcgttttca 300
tttaaagata ttcttgcgct ttaattacaa actgcaccga tgttggtggc gtcaaaatcg 360
ccgaggcgtt ccctgaaggc cggggcagcc cacatggatg tgggctgagg gcgcgtttta 420
caggg 425

<210> 32
<211> 381
<212> DNA
<213> Escherichia coli

<400> 32
aaaagagaat aaggtaaacc atccagctta atagctgatt aataactttt aacgcagctt 60
gttggcgaag cggaatgagg taacccacaa tcaggggaac cagaatgatt aacagcccag 120
aaaacatgaa aacccagtcc ttgcaaagat gaagtcgaaa tgcgcgatga cacactactg 180
aaagcggaag gacgagtaaa gttgcaatta aaaggaaatg ttatgcataa ggagcagtag 240
agtattcgtt ttcatttaaa gatattcttg cgctttaatt acaaactgca ccgatgttgg 300
tggcgtcaaa atcgccgagg cgttccctga aggccggggc agcccacatg gatgtgggct 360
gagggcgcgt tttacaggga t 381

<210> 33

<211> 329
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)...(329)
<223> n = A,T,C or G

<400> 33
gtcgatgttg tgcccgaact tgccatgctc cagtcctttt cttctgggcc gactttccgg 60
gaggtgctta aacgaaaaac cccgccccggt ttgcgccggc ggggttttgg aatcgtgtgt 120
tgttccagtc cctacggcgc attgccgacg accaccacca cacgcacgac gaccactgcg 180
gcggatgggc gcagttggta gtagttttgc gttgagcatg gaagtagtca ttggggacct 240
tgttggtttt gtgtttaaca atatttatac aagcacagct ttacagggga gacaatggaa 300
aatttttcag caagggaaaa ttgaggggn 329

<210> 34
<211> 442
<212> DNA
<213> Escherichia coli

<400> 34
gatgttgtgc ccgaacttgc catgctccag tccttttctt ctgggccgac tttccgggag 60
gtgcttaaac gaaaaacccc gccccggtttg cgccggcggg gttttggaat cgtgtgttgt 120
tccagtccct acggcgcatt gccgacgacc accaccacac gcacgacgac cactgcggcg 180
gatggcgcag ttggtagtag ttttgcgttg agcatggaag tagtcattgg ggaccttgtt 240
ggttttgtgt ttaacaatat ttatacaagc acagctttac aggggagaca atggaaaatt 300
tttcagcaag ggaaaattga ggggttgatc acgttttgta ctgaattgca gataacaaaa 360
aaccccgccg gagcgaggtt tcgtcagtcg cctgcggctg gtaaccgcaa agcacactgt 420
attatgtcaa cactgaaagt at 442

<210> 35
<211> 272
<212> DNA
<213> Escherichia coli

<400> 35
cgattatagg gctgaccaaa tacatggata gataatatgc catcccccca ggcagtccac 60
gtaaaaagag catatacaag gagaatagaa cacctacagc tgtaataaga taaacataat 120
gttttactct attaacacct acgatgaatg cataagagat gataaccata aagacaattt 180
gcttaccagt attcagtata attgcaaaga tgaatacaat aaaaacaagt aaagtaaatg 240
ttttacttac ctttgtatta gtaaattttt ta 272

<210> 36
<211> 255
<212> DNA
<213> Escherichia coli

<400> 36
tgcgattacg gcatgcttat gatcaggaat gttaatgcct gctatacggg ccactatgca 60
ctcctactat ttaatatgta cgttccatgc tgaaaagccc gttttcagga tactcaaatg 120
gaaacgcaca gacatacaaa agattggctg gctaatctag ccagctcaac ccaactttgc 180
aagaaaaata tgcgaaaaaa tcagccttgg cgctgtttat gcttcggctc ggcactgcaa 240
atcacacgga tgaca 255

<210> 37
<211> 389
<212> DNA
<213> Escherichia coli

<400> 37
aacgtaaaac tttacgtaaa ttaacatggt taacatttat gccactattg tttgtaaatt 60
catatttcgt aatgcttctg aatttttcgt gtgatggttt taaatactat ggtgtttact 120
cttgagggga cggcctattt ataaaatacg gacatttcaa taaatgcccg tataaacaga 180
gtatgattct ggctggtcgt tgagtatcaa tgttggaccg aatgtgaacg aataaataat 240
tcgtttattt taccacccat tctcttttcg attcgttgct tgtacgcgta gatgctacgg 300
atatccatat ccagaacttc agcaatctga tatggattat aaccagaaac catatacccc 360
attatggtga attccatctg agttaatgt 389

<210> 38
<211> 178
<212> DNA
<213> Escherichia coli

<400> 38
gtggttagga tctttatagt tacgctcggc ggcatccatc gggatacctg cggcattttc 60
tttggcaaaa ccgatttcag aattgtgttt gtttggatga acctgaatgg agagtggctg 120
tgctgcgcat aatactttga acaggaaagg cagttcgcca aagcgtttgg caacggcc 178

<210> 39
<211> 191
<212> DNA
<213> Escherichia coli

<400> 39
tcgcaaggaa aggcgtcagc gcaaaaacca gctccggctt gtggttagga tctttatagt 60
tacgctcggc ggcatccatc gggatacctg cggcattttc tttggcaaaa ccgatttcag 120
aattgtgttt gtttggatga acctgaatgg agagtggctg tgctgcgcat aatactttga 180
acaggaaagg c 191

<210> 40
<211> 158
<212> DNA
<213> Escherichia coli

<400> 40
ccggcttgtg gttaggatct ttatagttac gctcggcggc atccatcggg atacctgcgg 60
cattttcttt ggcaaaaccg atttcagaat tgtgtttgtt tggatgaacc tgaatggaga 120
gtggctgtgc tgcgcataat actttgaaca gaaaaggc 158

<210> 41
<211> 184
<212> DNA
<213> Escherichia coli

<400> 41
tcgaagaatt taacggaggg taaaaaaacc gacgcacact ggcgtcggct ctggcaggat 60
gtttcgtaat tagatagcca ccggcgcttt aatgcccgga tgcggatcgt agccttcaat 120
ctcaaagtct tcgaaacggt agtcgaagat ggattcgggt ttacgtttga taatcaactt 180
cggc 184

<210> 42
<211> 122
<212> DNA
<213> Escherichia coli

<400> 42
gtagaagttc aaattacgaa cctgaatttt actcggggca gtttcaacca tactcattgc 60
aatcttttcc tcatcgtgcc tgatgcactt cgctaatcag gcttaccgta ggccggatca 120
gg 122

<210> 43
<211> 163
<212> DNA
<213> Escherichia coli

<400> 43
taaattccgc gtatagcgtg ggcgctttgg ctccggtcgc ttcgacgagg tccgcaagag 60
aagtcgcttc ataaccgtgt tgccagaaca atttcatggc cttatcaagc gcggcatccc 120
tgtcgaacac ttttgggcgg ccacggcttt tttttacaca ttg 163

<210> 44
<211> 161
<212> DNA
<213> Escherichia coli

<400> 44
gcaaatgttt tcgtaattca ggagaaatgg caaataaagc attaaaaatt tgaatgcttt 60
gtgtaataaa aaagcagaca ggcgacggag tgaccactcc gtcgctttac aaagagagga 120
aaatcatagg ttgccggtgt agtgccagcg taaataacgc a 161

<210> 45
<211> 251
<212> DNA
<213> Escherichia coli

<400> 45
ggctgtaacg gtaaagccct caccgaagcg agggcttgaa ggagaagggt tatgatgcga 60
cttgtcatca tactgattgt actgttactc ataagtttca gcgcttatta acagtcagtc 120
tcaggggagg agcaatcctc ccttaccctt actcactaaa ttaggtcaaa gaatcaacga 180
tgtcaatcag ggcgatgcgg ttgtatcgcc cttaccactc ccagactttc gacggtgtaa 240
ccaccgcagg a 251

<210> 46
<211> 286
<212> DNA
<213> Escherichia coli

<400> 46
caccgaagcg agggcttgaa ggagaagggt tatgatgcga cttgtcatca tactgattgt 60
actgttactc ataagtttca gcgcttatta acagtcagtc tcaggggagg agcaatcctc 120
ccttaccctt actcactaaa ttaggtcaaa gaatcaacga tgtcaatcag ggcgatgcgg 180
gtgtatcgcc cttaccactc ccagactttc gacggtgtaa ccaccgcagg aagagggata 240
tcccactctt caacggggag tttttccacc aactgacaat catgcg 286

<210> 47
<211> 180
<212> DNA
<213> Escherichia coli

<400> 47
ccggggctga cgtgggcgat aatcgggtcg ccaggaatag ggcggcagca tttcgcaaag 60
gtgatcagca cgccatcggc acctttaatg ggcagatgtc cgtggctttg ggttgccggt 120
ggaatggagg cgtccccatg ttgcagattt ttcgcgacca ccacgctcat tgcgttacca 180

<210> 48
<211> 254
<212> DNA
<213> Escherichia coli

<400> 48

tgtcgagcat acgcggcagc gccgactcgg aagatgaagt ccccagtaca atcagcagtt 60
cttcacggat gtagcggata aatttgaaga tactgaaacc agtcgcttta gcgattgaac 120
ccaataccag caccacaaac aggatacagg taatgtagaa acagataatc agctgcccca 180
gttgcaccag tgtgccgacg ccgtatttac cgatggtaaa cgccattgcc ccgaacgcac 240
caataggtgc caga 254

<210> 49
<211> 300
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)...(300)
<223> n = A,T,C or G

<400> 49
agctgatcga tggtgtttgc caacgcacac gccaccgcgg cggcggcaaa caccancgta 60
ccgccgagca ccaccggctt acgcccgaag ctgtctgcca tcggcccgta gattaactgc 120
cccaacgcaa agcccagaat ataagtactg agggtcatct gcgtactgcc cgncggtacg 180
ccaaactgcg ctgaaattac cggtagcgcg ggcagataca tatcaatcga cagcggcatc 240
aacatggcca gcaggccaag gataaaaaca atagcaaacg acgaatgctg tcgggtggtc 300

<210> 50
<211> 207
<212> DNA
<213> Escherichia coli

<400> 50
caacatctca tgtctggatt tgtagatata cgtggaattc caattggtat cttttgtagg 60
attcaacata tcatgtaagc gctgggagct taccaccata actggcccat catcgactga 120
ggagaacagg ataaagatat tatcgtttcg tgttaattca ttcgtgttaa aggttcttaa 180
attatcttca gtttctataa atatagg 207

<210> 51
<211> 213
<212> DNA
<213> Escherichia coli

<400> 51
ttctgtaatc agaaaaagaa gaactggaat tttaataaat attatttctc tgggaacggg 60
ggggtataat caatatgacg agttgaatat tatatttttt ataagaatta tatgagatta 120
aagaaatctg ccgtaaagac agatttcttt aaaagataat tagagatttg cgacgttatg 180
ataaactttt tgtacatcgt cgtcatcttc aag 213

<210> 52
<211> 381
<212> DNA
<213> Escherichia coli

<400> 52
tctgttagtg tattatccac tgcggccctt tccgccgtct cgcaaacggg cgctggcttt 60
aggaaaggat gttccgtggc cgtaaatgca ggtgtttcac agcgcttgct atcgcggcaa 120
tatcgccagt ggtgctgtcg tgatgcggtc ttcgcatgga ccgcacaatg aagatacggt 180
gcttttgtat cgtacttatt gtttctggtg cgctgttaac cgaggtaaat aataaccgga 240
gtctctccgg cgacaattta ctggtggtta acaaccttca gagcagcaag taagcccgaa 300
tgccgccctt tgggcggcat attttagatt atccgattct gtttaaagtc acgcaaaaaa 360
ccacccccagc gacgttcata g 381

<210> 53

<211> 154
<212> DNA
<213> Escherichia coli

<400> 53
aaatagaccg gcatactttc gtgcggcagc ataatcaaca gtacgagcgc agcggtacta 60
atgataaagg tgtagcatag atatttactt ggcgatttca ggatcttatc tgaaatcatg 120
ccaccaatcg gtccaccaat cattttcaga cagt 154

<210> 54
<211> 191
<212> DNA
<213> Escherichia coli

<400> 54
tccttagtca ggtactgacg tactttaaaa tcgctgtcca ggttgtcagc gaattctttg 60
gtgttcgcaa accaggtaga gttccatggt tttacaatac ccaggcgaat accattagga 120
tgtactttct gacccattgc tagtctccag agtctcagcg atcggacaca accacagtga 180
tgtggctggt g 191

<210> 55
<211> 190
<212> DNA
<213> Escherichia coli

<400> 55
gctggtgccc gttttccatg ccagtggggc gacgcgcggc aaggcactat ccggcaaggg 60
ttgcgcttca tcagccataa tccggcgaat gatccacgcc gcccccgacg acattaaagg 120
ccgttcaagc agcggatcgt caggctgtaa gcgcaatttg cctgccttgc cgtggcgagc 180
aaacgcggta 190

<210> 56
<211> 402
<212> DNA
<213> Escherichia coli

<400> 56
aaaaaatgaa attcctcttt gacgggccaa tagcgatatt ggccattttt ttagcgcaac 60
atttgcggca aattcccttc tccatacagg tgtagtgcac cgaccgcgac cacatatcgc 120
cccggcggca tggcgcgtaa tttatcccgc caggcgagat ttcgctgatg catcagcaca 180
tcgtacagcg actgactgaa cgtattgggc agcgttatat cattattttg cggcggtgca 240
ttcagccacc agctcatcat ttgttgcagc aaccgtgcgt tggtatgcca gtgggtcagc 300
gtatcgtcca gcagcgccag tcctttgtca gggagctgga gcaacatggc aatctggttt 360
tcagcccctt ccagttcaat cacgggttta tgttgttgct tc 402

<210> 57
<211> 595
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)...(595)
<223> n = A,T,C or G

<400> 57
aattaatcag agcaacggta aaacaatgaa agtgtaaaaa acacttttgc gccaattatg 60
gagaaaaaaa gaaaatttga tggagagtga tgagagaata ttacaacacg atgatttgc 120
agagattatg aagaactata ccggatgact ggtgataaat aaagcaaata accaggatta 180
atctgtatta atttataaga aagcaactta atacccgcag aatgatttct gcgggtaagt 240
attagcttat tttttcgagc attaatcccg cgcgtaatcc caacgctacc aacggattag 300

ggaataacac atactctaca tcatgggtta cggtaaaacg ttcctctccg tcctgcgcca 360
gcaatgttcc tttctcaaac ggcataaaat tcagcgtgtc acttgccata tgcatttcga 420
aggacggcga gtggcgagta atttgcgaaa ccacccgata acggagcggc ggtgttctca 480
cgataccgac actctcacca gatagcagcg cagcaattgc gctggcagtt actgcaaact 540
ggcgaagatc gttttgccca aagggcaacg ctttgccaag ttncagccgt acang 595

<210> 58
<211> 250
<212> DNA
<213> Escherichia coli

<400> 58
gaaactcagt ggaatggggg agccgcaata gcgaacatgt tccatgccgc caataatctg 60
tgcgcggtga atatgcccga gcgcgatgta gtcggctggt ggaaagtttt gtgccggaaa 120
cgcgtccagc gtgccaatat aaatgtcacg cacggcgtca cttttactgg cccccacggt 180
cgttaaatgt cccgtggcga tgatgggcag aggctgatcg ccgcgcagtt tgcaggcatc 240
ggcatagtgt 250

<210> 59
<211> 236
<212> DNA
<213> Escherichia coli

<400> 59
aaaggttcgg caaaagccat aacctttcg tcagctttgg ccggataagg gacacattgc 60
gacgtttcga ccattttata acgggtaaaa ccaccatcaa catgagggaa atacatggca 120
ctgccaaaaa aacgcatatc tgtacactga ttctcgttat gttcaatgca gtatttgcag 180
tgaccgcacg gtttagacgg attaatggct accgtttgcc cttcatgtaa ttctga 236

<210> 60
<211> 92
<212> DNA
<213> Escherichia coli

<400> 60
gaagagatgt tcaggttttc gttatcggca atggtatcga acttgtattt ctcatacttc 60
tcgtcaggcg tggagtacgc cgcgccacgg aa 92

<210> 61
<211> 62
<212> DNA
<213> Escherichia coli

<400> 61
tgtcgacatt cagcacttcc gggtattcgt cgcgcagggc aaatgtacag gttgaggagg 60
tg 62

<210> 62
<211> 72
<212> DNA
<213> Escherichia coli

<400> 62
acttatcaaa ccatttttcc gcttcaccgg aggtctgcac ctgagcgatg gtgtcatcca 60
tcagcttttt ga 72

<210> 63
<211> 66
<212> DNA
<213> Escherichia coli

<400> 63
aaacgcatgg cgcgtaaagg ttcaatcagt ccgatttcag cggtgtcgaa ctcgctaaat 60
tcttca 66

<210> 64
<211> 143
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)...(143)
<223> n = A,T,C or G

<400> 64
tataggggnc tatttcatta acccgacact ggattaattc aacaggaata ttatgattct 60
gacacaactc tttaatttct tccgccgcca tcgtagaggt cgcaaccgcg cctccgcaag 120
cgacaataat cttgcgtttc ata 143

<210> 65
<211> 210
<212> DNA
<213> Escherichia coli

<400> 65
gaggaaccgt gtgacacaca aagggtgatt attttccatt gcggcgtata atcctggctc 60
accatcatgt attttcatgg ctaaaatact catcaccagt tttttgttta aaagctcttg 120
tgctattgga gtggttaatg cattcagaaa gatatcgatt atttcatcga atccatacaa 180
catggagata taaagtccag gtgtttcgtt 210

<210> 66
<211> 118
<212> DNA
<213> Escherichia coli

<400> 66
agagctgtga aggagaaaat acgaccacct ttaacggttt tagatacgcg gtttaccgcg 60
atcagctttt cctgcagttc gccagcttgt ttttcgatgt gagccatctt acacctct 118

<210> 67
<211> 531
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)...(531)
<223> n = A,T,C or G

<400> 67
cctttcaacg cagtagaggg gaatgncncc tacggcatcg gcaccgcgct caatttctgc 60
cagcatatag cccattccgt cttgatcggg cggaacatag gtgtcggcat cacttaaaaa 120
aacctggtcg caggtggcgt aattgaggcc attcatcagc gcaccacctt taccggtatt 180
tttttgcgtg acggcaacaa agcggtcgcc ccatttgcgt ttgacttccg ccatgaccgc 240
ttcggtattg tccgtggagc cgtcgttaac gcaaattacc cggcaaaaat aangggttcg 300
cagtaantta tccagtgact ggcgccagac acgggccttt tttntncgca gntttaatgg 360
cgtcnaatac agcctttttt tttantgggt ttaccntnta nccgtnttta aaaacccata 420
gcaaccattg ttttngacct tcanaatnaa aaattcnggg ttttaaaagc gcgttttccg 480
gcaacttgnt ggacagggca aaaaattcct gcctgggtcg ccgnatctgn g 531

<210> 68

<211> 102
<212> DNA
<213> Escherichia coli

<400> 68
agcgccggta acgcccgtta aatgttctcg ttcggcttcc cgatcgttat tcaacaggct 60
gcgtaccgta gcggcatatt tcttccctgc ttcatcgcca tc 102

<210> 69
<211> 167
<212> DNA
<213> Escherichia coli

<400> 69
aaaactcacc ttttttgttg ttatccctca attaccacaa gtaaaaggag ataaaccgac 60
aatgacctag acgtaattat gtaaaaatac tgatgttcgt cactgactat ttccgatatt 120
cagctgttga aaaatcaaaa aactggaaaa ataattttaa atatcat 167

<210> 70
<211> 83
<212> DNA
<213> Escherichia coli

<400> 70
gccacgccca gcataaacag cgggatcgag ccttcaatcg gtacgcccag tacacctttc 60
accatcagca ccagcgataa tcc 83

<210> 71
<211> 103
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)...(103)
<223> n = A,T,C or G

<400> 71
gccnggcggg aaccgacatg ttcccactct acgccggctg gcgcggtttt catttttgcg 60
cacctgatag ccgcacgatt gcagatcaac aacgcctttt ttg 103

<210> 72
<211> 121
<212> DNA
<213> Escherichia coli

<400> 72
tagcttcttc cacattgtcg aactgaacgt ccagacggtg tttagaaacg ttgatgctcg 60
ggattttttc cgcacccagg caaacgcgcg ccatggctgc aacggtttct gacggatact 120
g 121

<210> 73
<211> 163
<212> DNA
<213> Escherichia coli

<400> 73
tagttattgg ctgttttgag aatgtaatct tccaactgct tataactctt cagatgcaat 60
aatccgagaa tcacatgcaa tttattcata aattcgtggg atcgttcacg aagtgcgtca 120
gcatagttga ccagaccgtc gagtcgctgc atcagtttac gta 163

<210> 74
<211> 106
<212> DNA
<213> Escherichia coli

<400> 74
aaaaccatga ggttattatg gccgatttga ggagggaaag agtaagagca gtttgttaaa 60
tgtacaacga cgattctccc accgggcgcg ttttaaagcg acggtg 106

<210> 75
<211> 319
<212> DNA
<213> Escherichia coli

<400> 75
ctggagattg agtagatatt cttgttcaga atgtatcagc ccgatggttc tacgattctt 60
aagccacgaa gagttcagat agtacaacgg catgtctctt ttgactatct ggcaacccgg 120
cagtgtgttc tctcacgcat cacaaaagca gcaggcataa aaaaacccgc ttgcgcggct 180
ttttcacaaa gcttcagcaa attggcgatt aagccagttt gttgatctgt gcagtcaggt 240
tagccttatg acgtgcagct ttgtttttgt ggatcaaacc tttagcagcc tgacggtcca 300
cgatcggttg catttcgtt 319

<210> 76
<211> 237
<212> DNA
<213> Escherichia coli

<400> 76
cgaaccaacg acccccacca tgtcaaggtg gtgctctaac caactgagct atgaacgcaa 60
cgttgtaggt gacaacgggg acgaatatta gcggcagagt gggaaggtgg caagaggcaa 120
aacgtaattt tctgcgctat ttcgaccgtt tgcagagctt ttaagcaaat tggctatatt 180
ttgttgattt gcaagggtga tttttattca ggatcgcatt tacatctgat acaaccc 237

<210> 77
<211> 241
<212> DNA
<213> Escherichia coli

<400> 77
tgtgacagag tggttatcgt taatcaggca aagagggaaa aacatattat ttaaaccatt 60
atagcgcata aaatatcctt ataattaaca agaaagaaaa ggcatattct ctgcattatc 120
atttctacg actgtcaaaa atcgctcatt ttttaatgag tttatttgtt taatattatg 180
ggaaaaggtg atgcatttgg gagaggaaga gtattccccg gtcagacgac cggggaaggg 240
g 241

<210> 78
<211> 89
<212> DNA
<213> Escherichia coli

<400> 78
cccggctggg ttttttcaag tttacgcagg ggtgcgggcg atcacttccg ggttgctggc 60
agcttcgcct ttcggcgaaa tgctgtggt 89

<210> 79
<211> 140
<212> DNA
<213> Escherichia coli

<400> 79
tcgcgaacac cgtagtggat gtagttaccc gcagcatctt cgttgattgc tttagaacca 60

gaccacaggg tcaggttaga cggcgccagg tcagcagaac cgccgaggaa ttccggcaac 120
agcggaccga acgcttcgat 140

<210> 80
<211> 189
<212> DNA
<213> Escherichia coli

<400> 80
gttagtaaag ccgatgaggt gagcagtcac ttcgccggac ggatagtaac ggcgagactc 60
ttcacgcaga tgaatccccg gcagtttcag tttttgatg tagtccgcca tgtcagggtt 120
cacctgacgc gccagataaa taaagcgccc tttcgggttg gcgttaatgc gggctgaaag 180
ctgatccag 189

<210> 81
<211> 347
<212> DNA
<213> Escherichia coli

<400> 81
gcagttcacc ggcacgcgct ttcttatata agcctttggg atcgcgggct tcgcaaatcg 60
ccagcggcgt atcgacaaac acttcgataa agcgcccttc tcctacgcgt tcgcgaacca 120
tctggcgttc ggcgcggtgt ggcgagataa atgcggtcag caccaccagt ccggcttcaa 180
ccatcaaatt cgccacttca ccgacgcgac ggatattctc tttacgatcg gcatcgctaa 240
aaccgagatc gctgcataat ccgtggcgaa cattgtcgcc atccagcaga tacgtactga 300
cgccgagttt atgtaacgcc tcctccagcg ccccggcgac cgttgat 347

<210> 82
<211> 67
<212> DNA
<213> Escherichia coli

<400> 82
ccctgacgat cagtgatagt cacgatggtg ttgttgaaag aagcatggat atgagccacg 60
ccgtcag 67

<210> 83
<211> 176
<212> DNA
<213> Escherichia coli

<400> 83
agcaaagcga gccgaacatg gtggcgaggc tgccgtaaag taaaaactgg gtcgccaact 60
ccgggtgatc catcacgtat ttcacgaagt agagcgtcgc cccgccgcgc accacgttgg 120
agcaggtcgc catcatcttg aacgcgcaca tgatgcgcca ctggctgttg cccagc 176

<210> 84
<211> 632
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)...(632)
<223> n = A,T,C or G

<400> 84
tnncannnnc nnnnnnnnnn nnnnnncncn ntncctagan tntcnangtc tncncncncn 60
tnnccnnncc cttcnnnncc nnnnanangn ncnncncccc cnnncccnnt tttgttttgg 120
aaantccgtt tcaggaanag cttctgaatc ccgancgntt gataactgnn ggccagagtc 180
ataatgcgca ccaataataa tcagcggncc atcggcagga ccataatcag caacaatgnn 240

nnggtagggg ccacccgtaa tggggaccgt cctgcgaggt aactctggca ccgctactga 300
caaagacctc ttttatatat tctgcgggac ctattcagat tatcaatatt gtcnggcact 360
acgtggatgc acaagtttgt gtaagataac gaactgnttt ttctaattgc tcaggacttt 420
gcgtgtcggg gccgttgacg gtaatgcatt gacccagggt tggtaaaaaa taatcatagg 480
gagtaaaaac acaaataata taataaaagc caagattatt tttttcatat gcaaaattat 540
taatggggtt gcgtttggtg aatcattatn ttgctatggg tttcgatatt gngattttaa 600
ttaaagatta agcangggat tataccaaaa ga 632

<210> 85
<211> 161
<212> DNA
<213> Escherichia coli

<400> 85
gcgctattca gcgtaattgg cagcagctgc ggatcgtcat caatcttctc tgacagcggc 60
tgatgtactt caaccagacg cgcaccgttt ggttcggcag agacttttat cggagtgttg 120
atgatattca ctttggtgcc tggggtgacc tggctaaaga g 161

<210> 86
<211> 188
<212> DNA
<213> Escherichia coli

<400> 86
caagggtgag catttgcagc gaatcagaaa gggcgcagaa cagtgaagac aacagaaaaa 60
cgaccagacc gcatttataa atacgtcgat agccaaacat atcgcccaga aacgaaaacg 120
agagcaggga gatgacaatg gcgatttgat aggcgttcac tacccagatg gaactggctg 180
gcgtggca 188

<210> 87
<211> 175
<212> DNA
<213> Escherichia coli

<400> 87
tgagcatttg cagcgaatca gaaagggcgc agaacagtga agacaacaga aaaacgacca 60
gaccgcattt ataaatacgt cgatagccaa acatatcgcc cagaaacgaa aacgagagca 120
gggagatgac aatggcgatt tgataggcgt tcactaccca gatggaactg gctgg 175

<210> 88
<211> 194
<212> DNA
<213> Escherichia coli

<400> 88
cgcaagggtg agcatttgca gcgaatcaga aagggcgcag aacagtgaag acaacagaaa 60
aacgaccaga ccgcatttat aaatacgtcg atagccaaac atatcgccca gaaacgaaaa 120
cgagagcagg gagatgacaa tggcgatttg ataggcgttc actacccaga tggaactggc 180
tggcgtggca tgaa 194

<210> 89
<211> 272
<212> DNA
<213> Escherichia coli

<400> 89
ccttgtataa ggaaagggtt atgatgaagc tcgtcatcat actggttgtg ttgttactgt 60
taagtttccc gacttactaa caactcatca gagggggag aaatcctccc ttacccttgt 120
tcctttactc taggttgaaa aaacaacagc gtcaataggc ctgccatgta cgaagcgaga 180
tctgtgaacc gctttccggt tagccttttt tatcctgttg gatcttcttg atgatgttgg 240
tcgtcgagca accgtcttca aagttgagca cc 272

<210> 90
<211> 245
<212> DNA
<213> Escherichia coli

<400> 90
ccttgtataa ggaaagggtt atgatgaagc tcgtcatcat actggttgtg ttgttactgt 60
taagtttccc gacttactaa caactcatca gaggggggag aaatcctccc ttacccttgt 120
tcctttactc taggttgaaa aaacaacagc gtcaataggc ctgccatgta cgaagcgaga 180
tctgtgaacc gctttccggt tagccttttt tatcctgttg gatcttcttg atgatgttgg 240
tcgtc 245

<210> 91
<211> 203
<212> DNA
<213> Escherichia coli

<400> 91
taccaaaaaa agccacgtta tcttgttgat gcaaaagagt gaacgtggcg ttaaatgtaa 60
ccagttatat cagtagaaaa cctggttgtt gttaacagtc taaccggtca atttttttatg 120
attttttttga taaaaattaa attttatttg ctttaatcac caccagatga cgttcgccat 180
ccagggctgg aacctgaagt tta 203

<210> 92
<211> 189
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (1)...(189)
<223> n = A,T,C or G

<400> 92
ctaatgactt attttctgct taccaaaaaa agccacgtta tcttgttgat gcnaaagagt 60
gaacgtggcg ttaaatgtac cagttatatc agtagaaaac ctggttgntg ntaacagtct 120
aaccggcaat tttttatgat ttttttgata aaaattaaat tttatttgct ttaatcacca 180
ccagatgac 189

<210> 93
<211> 221
<212> DNA
<213> Escherichia coli

<400> 93
agcgcaacag cggaaaccag ggtagccagt tttttcatgt tcatattcaa gatgtcctgt 60
agtcgttatt actgcttttt gttgtctacc agcaccgcca gcaaaatcac caccgctttg 120
acgatcatct ggtaatagga ggaaacacct aacaaattca atccattatt aaggaagcca 180
agaattaatg cgccgatcaa cgtcccaaca atgcgacctt t 221

<210> 94
<211> 117
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(117)

<400> 94

```
atg aaa gtt cgt gct tcc gtc aag aaa tta tgc cgt aac tgc aaa atc    48
Met Lys Val Arg Ala Ser Val Lys Lys Leu Cys Arg Asn Cys Lys Ile
 1               5                  10                  15

gtt aag cgt gat ggt gtc atc cgt gtg att tgc agt gcc gag ccg aag    96
Val Lys Arg Asp Gly Val Ile Arg Val Ile Cys Ser Ala Glu Pro Lys
            20                  25                  30

cat aaa cag cgc caa ggc tga                                        117
His Lys Gln Arg Gln Gly  *
        35


<210> 95
<211> 1332
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1332)

<400> 95
atg gct aaa caa ccg gga tta gat ttt caa agt gcc aaa ggt ggc tta    48
Met Ala Lys Gln Pro Gly Leu Asp Phe Gln Ser Ala Lys Gly Gly Leu
 1               5                  10                  15

ggc gag ctg aaa cgc aga ctg ctg ttt gtt atc ggt gcg ctg att gtg    96
Gly Glu Leu Lys Arg Arg Leu Leu Phe Val Ile Gly Ala Leu Ile Val
            20                  25                  30

ttc cgt att ggc tct ttt att ccg atc cct ggt att gat gcc gct gta    144
Phe Arg Ile Gly Ser Phe Ile Pro Ile Pro Gly Ile Asp Ala Ala Val
        35                  40                  45

ctt gcc aaa ctg ctt gag caa cag cga ggc acc atc att gag atg ttt    192
Leu Ala Lys Leu Leu Glu Gln Gln Arg Gly Thr Ile Ile Glu Met Phe
    50                  55                  60

aac atg ttc tct ggt ggt gct ctc agc cgt gct tct atc ttt gct ctg    240
Asn Met Phe Ser Gly Gly Ala Leu Ser Arg Ala Ser Ile Phe Ala Leu
65                  70                  75                  80

ggg atc atg ccg tat att tcg gcg tcg atc att atc cag ctg ctg acg    288
Gly Ile Met Pro Tyr Ile Ser Ala Ser Ile Ile Ile Gln Leu Leu Thr
                85                  90                  95

gtg gtt cac cca acg ttg gca gaa att aag aaa gaa ggg gag tct ggt    336
Val Val His Pro Thr Leu Ala Glu Ile Lys Lys Glu Gly Glu Ser Gly
            100                 105                 110

cgt cgt aag atc agc cag tac acc cgc tac ggt act ctg gtg ctg gca    384
Arg Arg Lys Ile Ser Gln Tyr Thr Arg Tyr Gly Thr Leu Val Leu Ala
        115                 120                 125

ata ttc cag tcg atc ggt att gct acc ggt ctg ccg aat atg cct ggt    432
Ile Phe Gln Ser Ile Gly Ile Ala Thr Gly Leu Pro Asn Met Pro Gly
    130                 135                 140

atg caa ggc ctg gtg att aac ccg ggc ttt gca ttc tac ttc acc gct    480
Met Gln Gly Leu Val Ile Asn Pro Gly Phe Ala Phe Tyr Phe Thr Ala
```

```
145                 150                 155                 160

gtt gta agt ctg gtc aca gga acc atg ttc ctg atg tgg ttg ggc gaa     528
Val Val Ser Leu Val Thr Gly Thr Met Phe Leu Met Trp Leu Gly Glu
                165                 170                 175

cag att act gaa cga ggt atc ggc aac ggt att tca atc att atc ttc     576
Gln Ile Thr Glu Arg Gly Ile Gly Asn Gly Ile Ser Ile Ile Ile Phe
            180                 185                 190

gcc ggt att gtc gcg gga ctc ccg cca gcc att gcc cat act atc gag     624
Ala Gly Ile Val Ala Gly Leu Pro Pro Ala Ile Ala His Thr Ile Glu
        195                 200                 205

caa gcg cgt caa ggc gac ctg cac ttc ctc gtg ttg ctg ttg gtt gca     672
Gln Ala Arg Gln Gly Asp Leu His Phe Leu Val Leu Leu Leu Val Ala
    210                 215                 220

gta tta gta ttt gca gtg acg ttc ttt gtt gta ttt gtt gag cgt ggt     720
Val Leu Val Phe Ala Val Thr Phe Phe Val Val Phe Val Glu Arg Gly
225                 230                 235                 240

caa cgc cgc att gtg gta aac tac gcg aaa cgt cag caa ggt cgt cgt     768
Gln Arg Arg Ile Val Val Asn Tyr Ala Lys Arg Gln Gln Gly Arg Arg
                245                 250                 255

gtc tat gct gca cag agc aca cat tta ccg ctg aaa gtg aat atg gcg     816
Val Tyr Ala Ala Gln Ser Thr His Leu Pro Leu Lys Val Asn Met Ala
            260                 265                 270

ggg gta atc ccg gca atc ttc gct tcc agt att att ctg ttc ccg gcg     864
Gly Val Ile Pro Ala Ile Phe Ala Ser Ser Ile Ile Leu Phe Pro Ala
        275                 280                 285

acc atc gcg tca tgg ttc ggg ggc ggt act ggt tgg aac tgg ctg aca     912
Thr Ile Ala Ser Trp Phe Gly Gly Gly Thr Gly Trp Asn Trp Leu Thr
    290                 295                 300

aca att tcg ctg tat ttg cag cct ggg caa ccg ctt tat gtg tta ctc     960
Thr Ile Ser Leu Tyr Leu Gln Pro Gly Gln Pro Leu Tyr Val Leu Leu
305                 310                 315                 320

tat gcg tct gca atc atc ttc ttc tgt ttc ttc tac acg gcg ttg gtt     1008
Tyr Ala Ser Ala Ile Ile Phe Phe Cys Phe Phe Tyr Thr Ala Leu Val
                325                 330                 335

ttc aac ccg cgt gaa aca gca gat aac ctg aag aag tcc ggt gca ttt     1056
Phe Asn Pro Arg Glu Thr Ala Asp Asn Leu Lys Lys Ser Gly Ala Phe
            340                 345                 350

gta cca gga att cgt ccg gga gag caa acg gcg aag tat atc gat aaa     1104
Val Pro Gly Ile Arg Pro Gly Glu Gln Thr Ala Lys Tyr Ile Asp Lys
        355                 360                 365

gta atg acc cgc ctg acc ctg gtt ggt gcg ctg tat att acc ttt atc     1152
Val Met Thr Arg Leu Thr Leu Val Gly Ala Leu Tyr Ile Thr Phe Ile
    370                 375                 380

tgc ctg atc ccg gag ttc atg cgt gat gca atg aaa gta ccg ttc tac     1200
Cys Leu Ile Pro Glu Phe Met Arg Asp Ala Met Lys Val Pro Phe Tyr
385                 390                 395                 400
```

```
ttc ggt ggg acc tca ctg ctt atc gtt gtt gtc gtg att atg gac ttt    1248
Phe Gly Gly Thr Ser Leu Leu Ile Val Val Val Val Ile Met Asp Phe
                405                 410                 415

atg gct caa gtg caa act ctg atg atg tcc agt cag tat gag tct gca    1296
Met Ala Gln Val Gln Thr Leu Met Met Ser Ser Gln Tyr Glu Ser Ala
            420                 425                 430

ttg aag aag gcg aac ctg aaa ggc tac ggc cga taa                    1332
Leu Lys Lys Ala Asn Leu Lys Gly Tyr Gly Arg  *
        435                 440


<210> 96
<211> 435
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(435)

<400> 96
atg cgt tta aat act ctg tct ccg gcc gaa ggc tcc aaa aag gcg ggt    48
Met Arg Leu Asn Thr Leu Ser Pro Ala Glu Gly Ser Lys Lys Ala Gly
 1               5                  10                  15

aaa cgc ctg ggt cgt ggt atc ggt tct ggc ctc ggt aaa acc ggt ggt    96
Lys Arg Leu Gly Arg Gly Ile Gly Ser Gly Leu Gly Lys Thr Gly Gly
            20                  25                  30

cgt ggt cac aaa ggt cag aag tct cgt tct ggc ggt ggc gta cgt cgc    144
Arg Gly His Lys Gly Gln Lys Ser Arg Ser Gly Gly Gly Val Arg Arg
        35                  40                  45

ggt ttc gag ggt ggt cag atg cct ctg tac cgt cgt ctg ccg aaa ttc    192
Gly Phe Glu Gly Gly Gln Met Pro Leu Tyr Arg Arg Leu Pro Lys Phe
    50                  55                  60

ggc ttc act tct cgt aaa gca gcg att aca gcc gaa att cgt ctg tct    240
Gly Phe Thr Ser Arg Lys Ala Ala Ile Thr Ala Glu Ile Arg Leu Ser
 65                  70                  75                  80

gac ctg gct aaa gta gaa ggc ggt gta gta gac ctg aac acg ctg aaa    288
Asp Leu Ala Lys Val Glu Gly Gly Val Val Asp Leu Asn Thr Leu Lys
                85                  90                  95

gcg gct aac att atc ggt atc cag atc gag ttc gcg aaa gtg atc ctg    336
Ala Ala Asn Ile Ile Gly Ile Gln Ile Glu Phe Ala Lys Val Ile Leu
            100                 105                 110

gct ggc gaa gta acg act ccg gta act gtt cgt ggc ctg cgt gtt act    384
Ala Gly Glu Val Thr Thr Pro Val Thr Val Arg Gly Leu Arg Val Thr
        115                 120                 125

aaa ggc gct cgt gct gct atc gaa gct gct ggc ggt aaa atc gag gaa    432
Lys Gly Ala Arg Ala Ala Ile Glu Ala Ala Gly Gly Lys Ile Glu Glu
    130                 135                 140

taa                                                                435
```

*

<210> 97
<211> 180
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(180)

<400> 97
atg gca aag act att aaa att act caa acc cgc agt gca atc ggt cgt 48
Met Ala Lys Thr Ile Lys Ile Thr Gln Thr Arg Ser Ala Ile Gly Arg
1 5 10 15

ctg ccg aaa cac aag gca acg ctg ctt ggc ctg ggt ctg cgt cgt att 96
Leu Pro Lys His Lys Ala Thr Leu Leu Gly Leu Gly Leu Arg Arg Ile
20 25 30

ggt cac acc gta gag cgc gag gat act cct gct att cgc ggt atg atc 144
Gly His Thr Val Glu Arg Glu Asp Thr Pro Ala Ile Arg Gly Met Ile
35 40 45

aac gcg gtt tcc ttc atg gtt aaa gtt gag gag taa 180
Asn Ala Val Ser Phe Met Val Lys Val Glu Glu *
50 55

<210> 98
<211> 504
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(504)

<400> 98
atg gct cac atc gaa aaa caa gct ggc gaa ctg cag gaa aag ctg atc 48
Met Ala His Ile Glu Lys Gln Ala Gly Glu Leu Gln Glu Lys Leu Ile
1 5 10 15

gcg gta aac cgc gta tct aaa acc gtt aaa ggt ggt cgt att ttc tcc 96
Ala Val Asn Arg Val Ser Lys Thr Val Lys Gly Gly Arg Ile Phe Ser
20 25 30

ttc aca gct ctg act gta gtt ggc gat ggt aac ggt cgc gtt ggt ttt 144
Phe Thr Ala Leu Thr Val Val Gly Asp Gly Asn Gly Arg Val Gly Phe
35 40 45

ggt tac ggt aaa gcg cgt gaa gtt cca gca gcg atc cag aaa gcg atg 192
Gly Tyr Gly Lys Ala Arg Glu Val Pro Ala Ala Ile Gln Lys Ala Met
50 55 60

gaa aaa gcc cgt cgc aat atg att aac gtc gcg ctg aat aac ggc act 240
Glu Lys Ala Arg Arg Asn Met Ile Asn Val Ala Leu Asn Asn Gly Thr
65 70 75 80

ctg caa cac cct gtt aaa ggt gtt cac acg ggt tct cgc gta ttc atg 288
Leu Gln His Pro Val Lys Gly Val His Thr Gly Ser Arg Val Phe Met
85 90 95

cag ccg gct tcc gaa ggt acc ggt atc atc gcc ggt ggt gca atg cgc 336
Gln Pro Ala Ser Glu Gly Thr Gly Ile Ile Ala Gly Gly Ala Met Arg
100 105 110

gcc gtt ctg gaa gtc gct ggg gtt cat aac gtt ctg gct aaa gcc tat 384
Ala Val Leu Glu Val Ala Gly Val His Asn Val Leu Ala Lys Ala Tyr
115 120 125

ggt tcc acc aac ccg atc aac gtg gtt cgt gca act att gat ggc ctg 432
Gly Ser Thr Asn Pro Ile Asn Val Val Arg Ala Thr Ile Asp Gly Leu
130 135 140

gaa aat atg aat tct cca gaa atg gtc gct gcc aag cgt ggt aaa tcc 480
Glu Asn Met Asn Ser Pro Glu Met Val Ala Ala Lys Arg Gly Lys Ser
145 150 155 160

gtt gaa gaa att ctg ggg aaa taa 504
Val Glu Glu Ile Leu Gly Lys *
165

<210> 99
<211> 354
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(354)

<400> 99
atg gat aag aaa tct gct cgt atc cgt cgt gcg acc cgc gca cgc cgc 48
Met Asp Lys Lys Ser Ala Arg Ile Arg Arg Ala Thr Arg Ala Arg Arg
1 5 10 15

aag ctc cag gag ctg ggc gca act cgc ctg gtg gta cat cgt acc ccg 96
Lys Leu Gln Glu Leu Gly Ala Thr Arg Leu Val Val His Arg Thr Pro
20 25 30

cgt cac att tac gca cag gta att gca ccg aac ggt tct gaa gtt ctg 144
Arg His Ile Tyr Ala Gln Val Ile Ala Pro Asn Gly Ser Glu Val Leu
35 40 45

gta gct gct tct act gta gaa aaa gct atc gct gaa caa ctg aag tac 192
Val Ala Ala Ser Thr Val Glu Lys Ala Ile Ala Glu Gln Leu Lys Tyr
50 55 60

acc ggt aac aaa gac gcg gct gca gct gtg ggt aaa gct gtc gct gaa 240
Thr Gly Asn Lys Asp Ala Ala Ala Ala Val Gly Lys Ala Val Ala Glu
65 70 75 80

cgc gct ctg gaa aaa ggc atc aaa gat gta tcc ttt gac cgt tcc ggg 288
Arg Ala Leu Glu Lys Gly Ile Lys Asp Val Ser Phe Asp Arg Ser Gly
85 90 95

ttc caa tat cat ggt cgt gtc cag gca ctg gca gat gct gcc cgt gaa 336
Phe Gln Tyr His Gly Arg Val Gln Ala Leu Ala Asp Ala Ala Arg Glu

100 105 110

gct ggc ctt cag ttc taa 354
Ala Gly Leu Gln Phe *
115

<210> 100
<211> 534
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(534)

<400> 100
atg tct cgt gtt gct aaa gca ccg gtc gtt gtt cct gcc ggc gtt gac 48
Met Ser Arg Val Ala Lys Ala Pro Val Val Val Pro Ala Gly Val Asp
1 5 10 15

gta aaa atc aac ggt cag gtt att acg atc aaa ggt aaa aac ggc gag 96
Val Lys Ile Asn Gly Gln Val Ile Thr Ile Lys Gly Lys Asn Gly Glu
20 25 30

ctg act cgt act ctc aac gat gct gtt gaa gtt aaa cat gca gat aat 144
Leu Thr Arg Thr Leu Asn Asp Ala Val Glu Val Lys His Ala Asp Asn
35 40 45

acc ctg acc ttc ggt ccg cgt gat ggt tac gca gac ggt tgg gca cag 192
Thr Leu Thr Phe Gly Pro Arg Asp Gly Tyr Ala Asp Gly Trp Ala Gln
50 55 60

gct ggt acc gcg cgt gcc ctg ctg aac tca atg gtt atc ggt gtt acc 240
Ala Gly Thr Ala Arg Ala Leu Leu Asn Ser Met Val Ile Gly Val Thr
65 70 75 80

gaa ggc ttc act aag aag ctg cag ctg gtt ggt gta ggt tac cgt gca 288
Glu Gly Phe Thr Lys Lys Leu Gln Leu Val Gly Val Gly Tyr Arg Ala
85 90 95

gcg gtt aaa ggc aat gtg att aac ctg tct ctg ggt ttc tct cat cct 336
Ala Val Lys Gly Asn Val Ile Asn Leu Ser Leu Gly Phe Ser His Pro
100 105 110

gtt gac cat cag ctg cct gcg ggt atc act gct gaa tgt ccg act cag 384
Val Asp His Gln Leu Pro Ala Gly Ile Thr Ala Glu Cys Pro Thr Gln
115 120 125

act gaa atc gtg ctg aaa ggc gct gat aag cag gtg atc ggc cag gtt 432
Thr Glu Ile Val Leu Lys Gly Ala Asp Lys Gln Val Ile Gly Gln Val
130 135 140

gca gcg gat ctg cgc gcc tac cgt cgt cct gag cct tat aaa ggc aag 480
Ala Ala Asp Leu Arg Ala Tyr Arg Arg Pro Glu Pro Tyr Lys Gly Lys
145 150 155 160

ggt gtt cgt tac gcc gac gaa gtc gtg cgt acc aaa gag gct aag aag 528
Gly Val Arg Tyr Ala Asp Glu Val Val Arg Thr Lys Glu Ala Lys Lys
165 170 175

aag taa 534
Lys *

<210> 101
<211> 393
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(393)

<400> 101
atg agc atg caa gat ccg atc gcg gat atg ctg acc cgt atc cgt aac 48
Met Ser Met Gln Asp Pro Ile Ala Asp Met Leu Thr Arg Ile Arg Asn
1 5 10 15

ggt cag gcc gcg aac aaa gct gcg gtc acc atg cct tcc tcc aag ctg 96
Gly Gln Ala Ala Asn Lys Ala Ala Val Thr Met Pro Ser Ser Lys Leu
20 25 30

aaa gtg gca atc gcc aac gtg ctg aag gaa gaa ggt ttt att gaa gat 144
Lys Val Ala Ile Ala Asn Val Leu Lys Glu Glu Gly Phe Ile Glu Asp
35 40 45

ttt aaa gtt gaa ggc gac acc aag cct gaa ctg gaa ctt act ctg aag 192
Phe Lys Val Glu Gly Asp Thr Lys Pro Glu Leu Glu Leu Thr Leu Lys
50 55 60

tat ttc cag ggc aaa gct gtt gta gaa agc att cag cgt gtc agc cgc 240
Tyr Phe Gln Gly Lys Ala Val Val Glu Ser Ile Gln Arg Val Ser Arg
65 70 75 80

cca ggt ctg cgc atc tat aaa cgt aaa gat gag ctg ccg aaa gtt atg 288
Pro Gly Leu Arg Ile Tyr Lys Arg Lys Asp Glu Leu Pro Lys Val Met
85 90 95

gcg ggt ctg ggt atc gca gtt gtt tct acc tct aaa ggt gtt atg act 336
Ala Gly Leu Gly Ile Ala Val Val Ser Thr Ser Lys Gly Val Met Thr
100 105 110

gat cgt gca gcg cgc cag gct ggt ctt ggt ggc gaa att atc tgc tac 384
Asp Arg Ala Ala Arg Gln Ala Gly Leu Gly Gly Glu Ile Ile Cys Tyr
115 120 125

gta gcc taa 393
Val Ala *
130

<210> 102
<211> 306
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(306)

<400> 102
atg gct aag caa tca atg aaa gca cgc gaa gta aaa cgc gta gct tta 48
Met Ala Lys Gln Ser Met Lys Ala Arg Glu Val Lys Arg Val Ala Leu
1 5 10 15

gct gat aaa tac ttc gcg aaa cgc gct gaa ctg aaa gcg atc atc tct 96
Ala Asp Lys Tyr Phe Ala Lys Arg Ala Glu Leu Lys Ala Ile Ile Ser
20 25 30

gat gtg aac gct tcc gac gaa gat cgt tgg aac gct gtt ctc aag ctg 144
Asp Val Asn Ala Ser Asp Glu Asp Arg Trp Asn Ala Val Leu Lys Leu
35 40 45

cag act ctg ccg cgt gat tcc agc ccg tct cgt cag cgt aac cgc tgc 192
Gln Thr Leu Pro Arg Asp Ser Ser Pro Ser Arg Gln Arg Asn Arg Cys
50 55 60

cgt caa aca ggt cgt ccg cat ggt ttc ctg cgg aag ttc ggg ttg agc 240
Arg Gln Thr Gly Arg Pro His Gly Phe Leu Arg Lys Phe Gly Leu Ser
65 70 75 80

cgt att aag gtc cgt gaa gcc gct atg cgc ggt gaa atc ccg ggt ctg 288
Arg Ile Lys Val Arg Glu Ala Ala Met Arg Gly Glu Ile Pro Gly Leu
85 90 95

aaa aag gct agc tgg taa 306
Lys Lys Ala Ser Trp *
100

<210> 103
<211> 540
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(540)

<400> 103
atg gcg aaa ctg cat gat tac tac aaa gac gaa gta gtt aaa aaa ctc 48
Met Ala Lys Leu His Asp Tyr Tyr Lys Asp Glu Val Val Lys Lys Leu
1 5 10 15

atg act gag ttt aac tac aat tct gtc atg caa gtc cct cgg gtc gag 96
Met Thr Glu Phe Asn Tyr Asn Ser Val Met Gln Val Pro Arg Val Glu
20 25 30

aag atc acc ctg aac atg ggt gtt ggt gaa gcg atc gct gac aaa aaa 144
Lys Ile Thr Leu Asn Met Gly Val Gly Glu Ala Ile Ala Asp Lys Lys
35 40 45

ctg ctg gat aac gca gca gca gac ctg gca gca atc tcc ggt caa aaa 192
Leu Leu Asp Asn Ala Ala Ala Asp Leu Ala Ala Ile Ser Gly Gln Lys
50 55 60

ccg ctg atc acc aaa gca cgc aaa tct gtt gca ggc ttc aaa atc cgt 240
Pro Leu Ile Thr Lys Ala Arg Lys Ser Val Ala Gly Phe Lys Ile Arg
65 70 75 80

cag ggc tat ccg atc ggc tgt aaa gta act ctg cgt ggc gaa cgc atg 288

```
Gln Gly Tyr Pro Ile Gly Cys Lys Val Thr Leu Arg Gly Glu Arg Met
                 85                  90                  95

tgg gag ttc ttt gag cgc ctg atc act att gct gta cct cgt atc cgt     336
Trp Glu Phe Phe Glu Arg Leu Ile Thr Ile Ala Val Pro Arg Ile Arg
            100                 105                 110

gac ttc cgt ggc ctg tcc gct aag tct ttc gac ggt cgt ggt aac tac     384
Asp Phe Arg Gly Leu Ser Ala Lys Ser Phe Asp Gly Arg Gly Asn Tyr
        115                 120                 125

agc atg ggt gtc cgt gag cag atc atc ttc cca gaa atc gac tac gat     432
Ser Met Gly Val Arg Glu Gln Ile Ile Phe Pro Glu Ile Asp Tyr Asp
    130                 135                 140

aaa gtc gac cgc gtt cgt ggt ttg gat att acc att acc act act gcg     480
Lys Val Asp Arg Val Arg Gly Leu Asp Ile Thr Ile Thr Thr Thr Ala
145                 150                 155                 160

aaa tct gac gaa gaa ggc cgc gct ctg ctg gct gcc ttt gac ttc ccg     528
Lys Ser Asp Glu Glu Gly Arg Ala Leu Leu Ala Ala Phe Asp Phe Pro
                165                 170                 175

ttc cgc aag taa                                                     540
Phe Arg Lys  *
```

<210> 104
<211> 315
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(315)

<400> 104

```
atg gca gcg aaa atc cgt cgt gat gac gaa gtt atc gtg tta acc ggt     48
Met Ala Ala Lys Ile Arg Arg Asp Asp Glu Val Ile Val Leu Thr Gly
 1               5                  10                  15

aaa gat aaa ggt aaa cgc ggt aaa gtt aag aat gtc ctg tct tcc ggc     96
Lys Asp Lys Gly Lys Arg Gly Lys Val Lys Asn Val Leu Ser Ser Gly
            20                  25                  30

aag gtc att gtt gaa ggt atc aac ctg gtt aag aaa cat cag aag ccg     144
Lys Val Ile Val Glu Gly Ile Asn Leu Val Lys Lys His Gln Lys Pro
        35                  40                  45

gtt ccg gcc ctg aac caa ccg ggt ggc atc gtt gaa aaa gaa gcc gct     192
Val Pro Ala Leu Asn Gln Pro Gly Gly Ile Val Glu Lys Glu Ala Ala
    50                  55                  60

att cag gtt tcc aac gta gca atc ttc aat gcg gca acc ggc aag gct     240
Ile Gln Val Ser Asn Val Ala Ile Phe Asn Ala Ala Thr Gly Lys Ala
 65                 70                  75                  80

gac cgt gta ggc ttt aga ttc gaa gac ggt aaa aaa gtc cgt ttc ttc     288
Asp Arg Val Gly Phe Arg Phe Glu Asp Gly Lys Lys Val Arg Phe Phe
                 85                  90                  95
```

```
aag tct aac agc gaa act atc aag taa                                 315
Lys Ser Asn Ser Glu Thr Ile Lys  *
            100


<210> 105
<211> 372
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(372)

<400> 105
atg atc caa gaa cag act atg ctg aac gtc gcc gac aac tcc ggt gca    48
Met Ile Gln Glu Gln Thr Met Leu Asn Val Ala Asp Asn Ser Gly Ala
 1               5                  10                  15

cgt cgc gta atg tgt atc aag gtt ctg ggt ggc tcg cac cgt cgc tac    96
Arg Arg Val Met Cys Ile Lys Val Leu Gly Gly Ser His Arg Arg Tyr
            20                  25                  30

gca ggc gta ggc gac atc atc aag atc acc atc aaa gaa gca att ccg   144
Ala Gly Val Gly Asp Ile Ile Lys Ile Thr Ile Lys Glu Ala Ile Pro
        35                  40                  45

cgt ggt aag gtc aaa aaa ggt gat gtg ctg aag gcg gta gtg gtg cgc   192
Arg Gly Lys Val Lys Lys Gly Asp Val Leu Lys Ala Val Val Val Arg
    50                  55                  60

acc aag aag ggt gtt cgt cgc ccg gac ggt tct gtc att cgc ttc gat   240
Thr Lys Lys Gly Val Arg Arg Pro Asp Gly Ser Val Ile Arg Phe Asp
 65                  70                  75                  80

ggt aat gct tgt gtt ctt ctg aac aac aac agc gag cag cct atc ggt   288
Gly Asn Ala Cys Val Leu Leu Asn Asn Asn Ser Glu Gln Pro Ile Gly
                85                  90                  95

acg cgt att ttt ggg ccg gta act cgt gag ctt cgt agt gag aag ttc   336
Thr Arg Ile Phe Gly Pro Val Thr Arg Glu Leu Arg Ser Glu Lys Phe
            100                 105                 110

atg aaa att atc tct ctg gca cca gaa gta ctc taa                   372
Met Lys Ile Ile Ser Leu Ala Pro Glu Val Leu  *
        115                 120


<210> 106
<211> 1458
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1458)

<400> 106
gtg gga att tat ttt act aat tca gat gat caa att tac ttt aaa agg    48
Met Gly Ile Tyr Phe Thr Asn Ser Asp Asp Gln Ile Tyr Phe Lys Arg
```

1 5 10 15

agt gaa ggg atg tct gat ata aac cat gca ggt tct gac ctt ata ttt 96
Ser Glu Gly Met Ser Asp Ile Asn His Ala Gly Ser Asp Leu Ile Phe
20 25 30

gaa ctg gag gat cgc cct ccc ttt cat cag gct ctc gtt ggt gcc att 144
Glu Leu Glu Asp Arg Pro Pro Phe His Gln Ala Leu Val Gly Ala Ile
35 40 45

acc cat ctg ttg gca att ttc gtt ccg atg gta acc ccc gcg tta atc 192
Thr His Leu Leu Ala Ile Phe Val Pro Met Val Thr Pro Ala Leu Ile
50 55 60

gtg ggt gcg gcc tta cag ctt tcc gct gaa aca act gcc tat ctt gtt 240
Val Gly Ala Ala Leu Gln Leu Ser Ala Glu Thr Thr Ala Tyr Leu Val
65 70 75 80

tct atg gcg atg atc gcc tct ggt att ggt acc tgg tta caa gta aac 288
Ser Met Ala Met Ile Ala Ser Gly Ile Gly Thr Trp Leu Gln Val Asn
85 90 95

cgc tac ggc atc gtc ggt tct ggc cta ctc tca att cag tca gtc aat 336
Arg Tyr Gly Ile Val Gly Ser Gly Leu Leu Ser Ile Gln Ser Val Asn
100 105 110

ttt tca ttt gtt acg gtc atg att gcg ctg ggc agc agc atg aaa agc 384
Phe Ser Phe Val Thr Val Met Ile Ala Leu Gly Ser Ser Met Lys Ser
115 120 125

gac ggt ttt cac gaa gag tta atc atg tcg tcg ctt ctc ggc gtc tcc 432
Asp Gly Phe His Glu Glu Leu Ile Met Ser Ser Leu Leu Gly Val Ser
130 135 140

ttc gtt ggc gca ttt ctg gtt gtc gga tct tca ttt atc ttg ccc tat 480
Phe Val Gly Ala Phe Leu Val Val Gly Ser Ser Phe Ile Leu Pro Tyr
145 150 155 160

tta cgt cgg gtt att acg cct acc gtc agc ggt att gtg gta ctg atg 528
Leu Arg Arg Val Ile Thr Pro Thr Val Ser Gly Ile Val Val Leu Met
165 170 175

atc ggc tta agc ctg att aaa gtc ggc att atc gat ttt ggt gga gga 576
Ile Gly Leu Ser Leu Ile Lys Val Gly Ile Ile Asp Phe Gly Gly Gly
180 185 190

ttt gca gcc aaa agc agc ggt acg ttc ggc aat tac gaa cat ctc ggc 624
Phe Ala Ala Lys Ser Ser Gly Thr Phe Gly Asn Tyr Glu His Leu Gly
195 200 205

gtt ggt tta ttg gtt tta att gtg gtg atc ggc ttt aac tgc tgt cgc 672
Val Gly Leu Leu Val Leu Ile Val Val Ile Gly Phe Asn Cys Cys Arg
210 215 220

agt ccg ttg cta cgc atg gga ggg atc gcc att ggg cta tgt gtc ggc 720
Ser Pro Leu Leu Arg Met Gly Gly Ile Ala Ile Gly Leu Cys Val Gly
225 230 235 240

tat atc gca tcg tta tgc ctg ggc atg gtg gat ttc agc agt atg cgc 768
Tyr Ile Ala Ser Leu Cys Leu Gly Met Val Asp Phe Ser Ser Met Arg
245 250 255

```
aat ttg ccg tta atc acc atc ccg cat ccg ttc aaa tac ggc ttt agt    816
Asn Leu Pro Leu Ile Thr Ile Pro His Pro Phe Lys Tyr Gly Phe Ser
            260                 265                 270

ttt agc ttc cat cag ttc ctg gtg gtt ggc acg att tat ctg ctt agc    864
Phe Ser Phe His Gln Phe Leu Val Val Gly Thr Ile Tyr Leu Leu Ser
        275                 280                 285

gtg ctg gaa gca gtc ggc gat atc acc gcc acg gca atg gtt tcc cgc    912
Val Leu Glu Ala Val Gly Asp Ile Thr Ala Thr Ala Met Val Ser Arg
    290                 295                 300

cgc ccc att cag ggg gaa gag tat cag tcc cgg ctg aaa ggc ggc gtg    960
Arg Pro Ile Gln Gly Glu Glu Tyr Gln Ser Arg Leu Lys Gly Gly Val
305                 310                 315                 320

ctg gca gat ggt ctg gtt tct gtt atc gcc tcc gct gtc ggt tca tta    1008
Leu Ala Asp Gly Leu Val Ser Val Ile Ala Ser Ala Val Gly Ser Leu
                325                 330                 335

cca tta acc acg ttt gcg caa aat aat ggg gtt att cag atg act ggc    1056
Pro Leu Thr Thr Phe Ala Gln Asn Asn Gly Val Ile Gln Met Thr Gly
            340                 345                 350

gtc gct tca cgt tat gtc ggg cga acc atc gcg gta atg ctg gtt atc    1104
Val Ala Ser Arg Tyr Val Gly Arg Thr Ile Ala Val Met Leu Val Ile
        355                 360                 365

ctc ggc tta ttt ccg atg att ggc ggc ttc ttc acg acc att ccc tcg    1152
Leu Gly Leu Phe Pro Met Ile Gly Gly Phe Phe Thr Thr Ile Pro Ser
    370                 375                 380

gca gtt ctg gga ggc gca atg acg ttg atg ttt tcc atg att gcc atc    1200
Ala Val Leu Gly Gly Ala Met Thr Leu Met Phe Ser Met Ile Ala Ile
385                 390                 395                 400

gca ggg att cgc atc atc atc acc aac ggt tta aag cgc cgt gaa aca    1248
Ala Gly Ile Arg Ile Ile Ile Thr Asn Gly Leu Lys Arg Arg Glu Thr
                405                 410                 415

ctt att gtc gcc act tct tta ggt tta ggg ctt ggc gtc tcc tac gat    1296
Leu Ile Val Ala Thr Ser Leu Gly Leu Gly Leu Gly Val Ser Tyr Asp
            420                 425                 430

ccc gaa att ttt aaa ata ttg cca gcc tct att tat gta tta gtt gaa    1344
Pro Glu Ile Phe Lys Ile Leu Pro Ala Ser Ile Tyr Val Leu Val Glu
        435                 440                 445

aac cct att tgt gct ggc ggg tta act gcg att tta tta aat att atc    1392
Asn Pro Ile Cys Ala Gly Gly Leu Thr Ala Ile Leu Leu Asn Ile Ile
    450                 455                 460

ctc cct ggt ggc tac cga cag gaa aac gtt ctg cct ggt att acc tca    1440
Leu Pro Gly Gly Tyr Arg Gln Glu Asn Val Leu Pro Gly Ile Thr Ser
465                 470                 475                 480

gcg gaa gag atg gat taa                                            1458
Ala Glu Glu Met Asp  *
                485
```

<210> 107
<211> 1320
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1320)

<400> 107
atg atg tca gga gaa cac acg tta aaa gcg gta cga ggc agt ttt att 48
Met Met Ser Gly Glu His Thr Leu Lys Ala Val Arg Gly Ser Phe Ile
1 5 10 15

gat gtc acc cgt acg atc gat aac ccg gaa gag att gcc tct gcg ctg 96
Asp Val Thr Arg Thr Ile Asp Asn Pro Glu Glu Ile Ala Ser Ala Leu
20 25 30

cgg ttt att gag gat ggt tta tta ctc att aaa cag gga aaa gtg gaa 144
Arg Phe Ile Glu Asp Gly Leu Leu Leu Ile Lys Gln Gly Lys Val Glu
35 40 45

tgg ttt ggc gaa tgg gaa aac gga aag cat caa att cct gac acc att 192
Trp Phe Gly Glu Trp Glu Asn Gly Lys His Gln Ile Pro Asp Thr Ile
50 55 60

cgc gtg cgc gac tat cgc ggc aaa ctg ata gta ccg ggc ttt gtc gat 240
Arg Val Arg Asp Tyr Arg Gly Lys Leu Ile Val Pro Gly Phe Val Asp
65 70 75 80

aca cat atc cat tat ccg caa agt gaa atg gtg ggg gcc tat ggt gag 288
Thr His Ile His Tyr Pro Gln Ser Glu Met Val Gly Ala Tyr Gly Glu
85 90 95

caa ttg ctg gag tgg ttg aat aaa cac acc ttc cct act gaa cgt cgt 336
Gln Leu Leu Glu Trp Leu Asn Lys His Thr Phe Pro Thr Glu Arg Arg
100 105 110

tat gag gat tta gag tac gcc cgc gaa atg tcg gcg ttc ttc atc aag 384
Tyr Glu Asp Leu Glu Tyr Ala Arg Glu Met Ser Ala Phe Phe Ile Lys
115 120 125

cag ctt tta cgt aac gga acc acc acg gcg ctg gtg ttt ggc act gtt 432
Gln Leu Leu Arg Asn Gly Thr Thr Thr Ala Leu Val Phe Gly Thr Val
130 135 140

cat ccg caa tct gtt gat gcg ctg ttt gaa gcc gcc agt cat atc aat 480
His Pro Gln Ser Val Asp Ala Leu Phe Glu Ala Ala Ser His Ile Asn
145 150 155 160

atg cgt atg att gcc ggt aag gtg atg atg gac cgc aac gca ccg gat 528
Met Arg Met Ile Ala Gly Lys Val Met Met Asp Arg Asn Ala Pro Asp
165 170 175

tat ctg ctc gac act gcc gaa agc agc tat cac caa agc aaa gaa ctg 576
Tyr Leu Leu Asp Thr Ala Glu Ser Ser Tyr His Gln Ser Lys Glu Leu
180 185 190

atc gaa cgc tgg cac aaa aat ggt cgt ctg cta tat gcg att acg cca 624
Ile Glu Arg Trp His Lys Asn Gly Arg Leu Leu Tyr Ala Ile Thr Pro

```
        195                 200                 205

cgc ttc gcc ccg acc tca tct cct gaa cag atg gcg atg gcg caa cgc      672
Arg Phe Ala Pro Thr Ser Ser Pro Glu Gln Met Ala Met Ala Gln Arg
    210                 215                 220

ctg aaa gaa gaa tat ccg gat acg tgg gta cat acc cat ctc tgt gaa      720
Leu Lys Glu Glu Tyr Pro Asp Thr Trp Val His Thr His Leu Cys Glu
225                 230                 235                 240

aac aaa gat gaa att gcc tgg gtg aaa tcg ctt tat cct gac cat gat      768
Asn Lys Asp Glu Ile Ala Trp Val Lys Ser Leu Tyr Pro Asp His Asp
                245                 250                 255

ggt tat ctg gat gtt tac cat cag tac ggc ctg acc ggt aaa aac tgt      816
Gly Tyr Leu Asp Val Tyr His Gln Tyr Gly Leu Thr Gly Lys Asn Cys
            260                 265                 270

gtc ttt gct cac tgc gtc cat ctc gaa gaa aaa gag tgg gat cgt ctc      864
Val Phe Ala His Cys Val His Leu Glu Glu Lys Glu Trp Asp Arg Leu
        275                 280                 285

agc gaa acc aaa tcc agc att gct ttc tgt ccg acc tcc aac ctt tac      912
Ser Glu Thr Lys Ser Ser Ile Ala Phe Cys Pro Thr Ser Asn Leu Tyr
    290                 295                 300

ctc ggc agc ggc tta ttc aac ttg aaa aaa gca tgg cag aag aaa gtt      960
Leu Gly Ser Gly Leu Phe Asn Leu Lys Lys Ala Trp Gln Lys Lys Val
305                 310                 315                 320

aaa gtg ggc atg gga acg gat atc ggt gcc gga acc act ttc aac atg      1008
Lys Val Gly Met Gly Thr Asp Ile Gly Ala Gly Thr Thr Phe Asn Met
                325                 330                 335

ctg caa acg ctg aac gaa gcc tac aaa gta ttg caa tta caa ggc tat      1056
Leu Gln Thr Leu Asn Glu Ala Tyr Lys Val Leu Gln Leu Gln Gly Tyr
            340                 345                 350

cgc ctc tcg gca tat gaa gcg ttt tac ctg gcc acg ctc ggc gga gcg      1104
Arg Leu Ser Ala Tyr Glu Ala Phe Tyr Leu Ala Thr Leu Gly Gly Ala
        355                 360                 365

aaa tct ctg ggc ctt gac gat ttg att ggc aac ttt tta cct ggc aaa      1152
Lys Ser Leu Gly Leu Asp Asp Leu Ile Gly Asn Phe Leu Pro Gly Lys
    370                 375                 380

gag gct gat ttc gtg gtg atg gaa ccc acc gcc act ccg cta cag cag      1200
Glu Ala Asp Phe Val Val Met Glu Pro Thr Ala Thr Pro Leu Gln Gln
385                 390                 395                 400

ctg cgc tat gac aac tct gtt tct tta gtc gac aaa ttg ttc gtg atg      1248
Leu Arg Tyr Asp Asn Ser Val Ser Leu Val Asp Lys Leu Phe Val Met
                405                 410                 415

atg acg ttg ggc gat gac cgt tcg atc tac cgc acc tac gtt gat ggt      1296
Met Thr Leu Gly Asp Asp Arg Ser Ile Tyr Arg Thr Tyr Val Asp Gly
            420                 425                 430

cgt ctg gtg tac gaa cgc aac taa                                      1320
Arg Leu Val Tyr Glu Arg Asn  *
        435
```

<210> 108
<211> 570
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(570)

<400> 108
atg tct gga gac atc cta caa aca ccg gac gca cca aag cca cag ggc 48
Met Ser Gly Asp Ile Leu Gln Thr Pro Asp Ala Pro Lys Pro Gln Gly
1 5 10 15

gcg ctg gat aat tat ttt aaa att acc gct cgt ggc agt acc gtt cgt 96
Ala Leu Asp Asn Tyr Phe Lys Ile Thr Ala Arg Gly Ser Thr Val Arg
20 25 30

cag gaa gta ctg gct ggc tta acg acc ttt ctg gcc atg gtt tat tcc 144
Gln Glu Val Leu Ala Gly Leu Thr Thr Phe Leu Ala Met Val Tyr Ser
35 40 45

gtt atc gtc gtt ccg gga atg ctg ggc aaa gca ggt ttt cct ccc gca 192
Val Ile Val Val Pro Gly Met Leu Gly Lys Ala Gly Phe Pro Pro Ala
50 55 60

gct gtg ttt gtt gcc acc tgt ctg gtc gcg ggc ttc ggc tcg ttg ctg 240
Ala Val Phe Val Ala Thr Cys Leu Val Ala Gly Phe Gly Ser Leu Leu
65 70 75 80

atg gga tta tgg gct aat ttg cca atg gcg att ggt tgc gcg att tcc 288
Met Gly Leu Trp Ala Asn Leu Pro Met Ala Ile Gly Cys Ala Ile Ser
85 90 95

ttg acg gcg ttt acc gca ttc agt ctg gta ctc ggg caa caa att agc 336
Leu Thr Ala Phe Thr Ala Phe Ser Leu Val Leu Gly Gln Gln Ile Ser
100 105 110

gtt cct gtc gca ctg ggc gcg gta ttt ctg atg ggc gtc atc ttc acc 384
Val Pro Val Ala Leu Gly Ala Val Phe Leu Met Gly Val Ile Phe Thr
115 120 125

gcc att tcc gta acc ggt gtg cgt acc tgg atc tta cgt aat ttg ccg 432
Ala Ile Ser Val Thr Gly Val Arg Thr Trp Ile Leu Arg Asn Leu Pro
130 135 140

atg ggt atc gct cac ggt aca ggt atc ggt atc ggg ctg ttt ctg ctg 480
Met Gly Ile Ala His Gly Thr Gly Ile Gly Ile Gly Leu Phe Leu Leu
145 150 155 160

ctg att gct gct aac ggt gtg ggt atg gtt atc aaa aac ccg att gaa 528
Leu Ile Ala Ala Asn Gly Val Gly Met Val Ile Lys Asn Pro Ile Glu
165 170 175

ggc ttg cag tgg cgc tcg gtg cgt tta cct cct tcc cgg tga 570
Gly Leu Gln Trp Arg Ser Val Arg Leu Pro Pro Ser Arg *
180 185

<210> 109
<211> 831
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(831)

<400> 109
gtg gcg ctc ggt gcg ttt acc tcc ttc ccg gtg atg atg agc ttg ctg 48
Met Ala Leu Gly Ala Phe Thr Ser Phe Pro Val Met Met Ser Leu Leu
1 5 10 15

ggg ctg gcg gtc atc ttc ggc ctg gag aag tgt cgc gta ccc ggc ggg 96
Gly Leu Ala Val Ile Phe Gly Leu Glu Lys Cys Arg Val Pro Gly Gly
20 25 30

atc ttg ttg gtg att att gca att tcg atc atc ggc tta atc ttt gac 144
Ile Leu Leu Val Ile Ile Ala Ile Ser Ile Ile Gly Leu Ile Phe Asp
35 40 45

cca gcg gtg aaa tac cac ggt ctg gtg gcg atg cca agc ctg act ggc 192
Pro Ala Val Lys Tyr His Gly Leu Val Ala Met Pro Ser Leu Thr Gly
50 55 60

gaa gat ggt aag tct ctg att ttc agc ctc gat att atg ggt gca ctc 240
Glu Asp Gly Lys Ser Leu Ile Phe Ser Leu Asp Ile Met Gly Ala Leu
65 70 75 80

cag cca act gta ctt ccg agt gta ctg gca ttg gtg atg acc gca gtg 288
Gln Pro Thr Val Leu Pro Ser Val Leu Ala Leu Val Met Thr Ala Val
85 90 95

ttc gac gct act ggc acc atc cgt gcc gtc gcc ggt cag gcg aat ttg 336
Phe Asp Ala Thr Gly Thr Ile Arg Ala Val Ala Gly Gln Ala Asn Leu
100 105 110

ttg gat aaa gac aac cag atc atc aac ggc ggc aaa gcc ctg acc agt 384
Leu Asp Lys Asp Asn Gln Ile Ile Asn Gly Gly Lys Ala Leu Thr Ser
115 120 125

gac tca gta agt tca ata ttc tcc ggc ctg gtg ggc gca gcg ccc gca 432
Asp Ser Val Ser Ser Ile Phe Ser Gly Leu Val Gly Ala Ala Pro Ala
130 135 140

gcg gtt tat atc gaa tca gcg gca gga acc gcc gcc ggg ggt aaa aca 480
Ala Val Tyr Ile Glu Ser Ala Ala Gly Thr Ala Ala Gly Gly Lys Thr
145 150 155 160

ggt tta acc gca acc gta gtg ggg gcg tta ttc ctg tta att ctg ttt 528
Gly Leu Thr Ala Thr Val Val Gly Ala Leu Phe Leu Leu Ile Leu Phe
165 170 175

tta tca ccg ctg tca ttt ttg atc cct ggt tac gcc act gca ccc gct 576
Leu Ser Pro Leu Ser Phe Leu Ile Pro Gly Tyr Ala Thr Ala Pro Ala
180 185 190

ctg atg tac gta ggt ttg ctg atg tta agt aac gtc tcg aag ctg gat 624
Leu Met Tyr Val Gly Leu Leu Met Leu Ser Asn Val Ser Lys Leu Asp

```
        195                 200                 205

ttc aat gat ttt att gac gct atg gct ggc ctg gtg tgt gcc gtg ttc    672
Phe Asn Asp Phe Ile Asp Ala Met Ala Gly Leu Val Cys Ala Val Phe
    210                 215                 220

atc gtt ctg act tgt aat atc gtt acc ggt att atg ctg ggc ttt gtg    720
Ile Val Leu Thr Cys Asn Ile Val Thr Gly Ile Met Leu Gly Phe Val
225                 230                 235                 240

aca ctg gtc gta ggc cgc gtc ttt gca cgc gaa tgg caa aag ctg aat    768
Thr Leu Val Val Gly Arg Val Phe Ala Arg Glu Trp Gln Lys Leu Asn
                245                 250                 255

att ggt acg gtg atc att act gcc gca ctg gtc gca ttt tac gcg ggt    816
Ile Gly Thr Val Ile Ile Thr Ala Ala Leu Val Ala Phe Tyr Ala Gly
            260                 265                 270

ggt tgg gca atc taa                                                831
Gly Trp Ala Ile  *
        275
```

<210> 110
<211> 1401
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1401)

<400> 110

```
atg aat agc gaa ggg ggg aaa ccg ggg aat gta ctg acc gtt aac ggc    48
Met Asn Ser Glu Gly Gly Lys Pro Gly Asn Val Leu Thr Val Asn Gly
 1               5                  10                  15

aac tat acc gga aac aat ggc ctg atg acg ttc aac gcg acg ctg ggc    96
Asn Tyr Thr Gly Asn Asn Gly Leu Met Thr Phe Asn Ala Thr Leu Gly
            20                  25                  30

ggc gat aat tcg ccc acc gat aag atg aac gtg aaa ggc gat acc caa    144
Gly Asp Asn Ser Pro Thr Asp Lys Met Asn Val Lys Gly Asp Thr Gln
        35                  40                  45

ggg aac act cgc gtt cgg gtt gat aac att ggc ggc gtc ggt gca caa    192
Gly Asn Thr Arg Val Arg Val Asp Asn Ile Gly Gly Val Gly Ala Gln
    50                  55                  60

acg gtc aac ggt att gaa ctc att gag gtt ggc ggt aat tct gca ggt    240
Thr Val Asn Gly Ile Glu Leu Ile Glu Val Gly Gly Asn Ser Ala Gly
 65                 70                  75                  80

aac ttc gcg ctg acc acc gga act gtc gaa gct ggg gct tac gtc tac    288
Asn Phe Ala Leu Thr Thr Gly Thr Val Glu Ala Gly Ala Tyr Val Tyr
                85                  90                  95

acg ctg gct aaa ggg aag ggg aat gac gag aaa aac tgg tat ctg acc    336
Thr Leu Ala Lys Gly Lys Gly Asn Asp Glu Lys Asn Trp Tyr Leu Thr
            100                 105                 110
```

```
agt aaa tgg gac ggc gta acg cca gcg gat aca ccc gat ccc atc aat   384
Ser Lys Trp Asp Gly Val Thr Pro Ala Asp Thr Pro Asp Pro Ile Asn
        115                 120                 125

aat ccc cct gtt gtg gat ccg gaa ggc cca tca gtt tat cgc ccg gag   432
Asn Pro Pro Val Val Asp Pro Glu Gly Pro Ser Val Tyr Arg Pro Glu
    130                 135                 140

gcc gga agc tat atc agc aac att gcc gca gcc aac tcg ctg ttt agc   480
Ala Gly Ser Tyr Ile Ser Asn Ile Ala Ala Ala Asn Ser Leu Phe Ser
145                 150                 155                 160

cat cgt tta cac gac cgt ctg ggt gag ccg cag tat aca gat tca ctg   528
His Arg Leu His Asp Arg Leu Gly Glu Pro Gln Tyr Thr Asp Ser Leu
                165                 170                 175

cat tct cag ggg tcg gca agc agt atg tgg atg cgt cat gtc gga ggg   576
His Ser Gln Gly Ser Ala Ser Ser Met Trp Met Arg His Val Gly Gly
            180                 185                 190

cac gaa cgt tca agg gcc ggt gac ggt cag cta aat act cag gct aac   624
His Glu Arg Ser Arg Ala Gly Asp Gly Gln Leu Asn Thr Gln Ala Asn
        195                 200                 205

cgc tat gta ttg cag cta ggc ggc gat ttg gcg cag tgg agt agc aac   672
Arg Tyr Val Leu Gln Leu Gly Gly Asp Leu Ala Gln Trp Ser Ser Asn
    210                 215                 220

gcg cag gat cgc tgg cat ctt ggc gtg atg gca ggc tac gcc aat cag   720
Ala Gln Asp Arg Trp His Leu Gly Val Met Ala Gly Tyr Ala Asn Gln
225                 230                 235                 240

cac agt aat act cag agt aat cgt gtg ggt tat aaa tcg gat ggg cgc   768
His Ser Asn Thr Gln Ser Asn Arg Val Gly Tyr Lys Ser Asp Gly Arg
                245                 250                 255

atc agc ggt tac agc gct ggg ctg tac gcg acc tgg tat cag aac gat   816
Ile Ser Gly Tyr Ser Ala Gly Leu Tyr Ala Thr Trp Tyr Gln Asn Asp
            260                 265                 270

gcg aat aag acc ggc gct tat gtt gac agc tgg gcg ctg tat aac tgg   864
Ala Asn Lys Thr Gly Ala Tyr Val Asp Ser Trp Ala Leu Tyr Asn Trp
        275                 280                 285

ttt gat aac agc gtc agt tcc gat aac cgt tct gct gac gac tat gat   912
Phe Asp Asn Ser Val Ser Ser Asp Asn Arg Ser Ala Asp Asp Tyr Asp
    290                 295                 300

tct cgc ggt gtg acg gcc tct gtt gag ggt ggg tat acc ttt gaa gcg   960
Ser Arg Gly Val Thr Ala Ser Val Glu Gly Gly Tyr Thr Phe Glu Ala
305                 310                 315                 320

gga aca ttt agc ggc agc gaa ggg acg ctg aat acc tgg tac gtc cag   1008
Gly Thr Phe Ser Gly Ser Glu Gly Thr Leu Asn Thr Trp Tyr Val Gln
                325                 330                 335

cca cag gcg caa atc acc tgg atg ggt gtg aaa gat tcc gac cat acc   1056
Pro Gln Ala Gln Ile Thr Trp Met Gly Val Lys Asp Ser Asp His Thr
            340                 345                 350

cgg aaa gac gga acg cgc att gaa acg gaa ggc gac gga aat gtg caa   1104
```

```
Arg Lys Asp Gly Thr Arg Ile Glu Thr Glu Gly Asp Gly Asn Val Gln
        355                 360                 365

acg cga ctt ggg gtg aaa acc tac ctg aac agc cat cac cag cgt gac    1152
Thr Arg Leu Gly Val Lys Thr Tyr Leu Asn Ser His His Gln Arg Asp
    370                 375                 380

gat ggt aaa cag cgt gag ttc cag cct tac att gaa gcg aac tgg atc    1200
Asp Gly Lys Gln Arg Glu Phe Gln Pro Tyr Ile Glu Ala Asn Trp Ile
385                 390                 395                 400

aac aat agc aaa gtc tac gcc gtg aag atg aat ggt caa acc gta ggc    1248
Asn Asn Ser Lys Val Tyr Ala Val Lys Met Asn Gly Gln Thr Val Gly
                405                 410                 415

cgt gaa ggt gcg cgt aat ctc ggt gaa gta cgt acc ggg gtt gag gcg    1296
Arg Glu Gly Ala Arg Asn Leu Gly Glu Val Arg Thr Gly Val Glu Ala
            420                 425                 430

aaa gta aat aac aac ctt agc ctg tgg ggg aat gtc ggt gtg caa cta    1344
Lys Val Asn Asn Asn Leu Ser Leu Trp Gly Asn Val Gly Val Gln Leu
        435                 440                 445

ggt gat aaa ggc tat agc gat act cag ggc atg ctg gga gtg aaa tat    1392
Gly Asp Lys Gly Tyr Ser Asp Thr Gln Gly Met Leu Gly Val Lys Tyr
    450                 455                 460

agc tgg taa                                                        1401
Ser Trp  *
465


<210> 111
<211> 3978
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(3978)

<400> 111
atg aat aga atc tat cgc gtg ata tgg aat tgc act cta cag gta ttt    48
Met Asn Arg Ile Tyr Arg Val Ile Trp Asn Cys Thr Leu Gln Val Phe
 1               5                  10                  15

cag gcc tgc tcg gaa tta act cgc agg gca ggt aaa aca tcg acg gtt    96
Gln Ala Cys Ser Glu Leu Thr Arg Arg Ala Gly Lys Thr Ser Thr Val
            20                  25                  30

aat ttg cgt aaa tcc tct gga ctg aca acg aaa ttc agt aga ttg acg    144
Asn Leu Arg Lys Ser Ser Gly Leu Thr Thr Lys Phe Ser Arg Leu Thr
        35                  40                  45

ctg ggt gtt ttg ctg gca cta agc ggt tca gcg tct ggt gca agt ctg    192
Leu Gly Val Leu Leu Ala Leu Ser Gly Ser Ala Ser Gly Ala Ser Leu
    50                  55                  60

gaa gtt gat aat gat cag att acc aat att gat act gat gtt gct tat    240
Glu Val Asp Asn Asp Gln Ile Thr Asn Ile Asp Thr Asp Val Ala Tyr
 65                  70                  75                  80
```

gat gcc tac ctg gtt ggc tgg tat ggc act gga gtg ctt aat att ttg 288
Asp Ala Tyr Leu Val Gly Trp Tyr Gly Thr Gly Val Leu Asn Ile Leu
85 90 95

gct ggc ggt aat gcc tcc tta acc act att act acc agc gtc att ggc 336
Ala Gly Gly Asn Ala Ser Leu Thr Thr Ile Thr Thr Ser Val Ile Gly
100 105 110

gct aat gag gac tca gag ggc acc gtt aat gtt ttg ggt ggc acc tgg 384
Ala Asn Glu Asp Ser Glu Gly Thr Val Asn Val Leu Gly Gly Thr Trp
115 120 125

cga ttg tat gat agc gga aat aat gca agg cct tta aat gtg ggt caa 432
Arg Leu Tyr Asp Ser Gly Asn Asn Ala Arg Pro Leu Asn Val Gly Gln
130 135 140

tcc gga acg ggg acg ctg aat att aaa cag aag ggt cac gtc gat gga 480
Ser Gly Thr Gly Thr Leu Asn Ile Lys Gln Lys Gly His Val Asp Gly
145 150 155 160

ggc tat tta aga tta ggt tct tcg aca gga ggc gtc ggg acg gtc aat 528
Gly Tyr Leu Arg Leu Gly Ser Ser Thr Gly Gly Val Gly Thr Val Asn
165 170 175

gtt gag gga gag gac tct gtt ctg acg acc gaa tta ttc gaa ata ggt 576
Val Glu Gly Glu Asp Ser Val Leu Thr Thr Glu Leu Phe Glu Ile Gly
180 185 190

agc tat ggc aca ggt tca tta aat att acg gat aag ggt tac gtc acg 624
Ser Tyr Gly Thr Gly Ser Leu Asn Ile Thr Asp Lys Gly Tyr Val Thr
195 200 205

agt tca atc gtc gcc att tta ggc tat cag gcg ggc agt aat ggc cag 672
Ser Ser Ile Val Ala Ile Leu Gly Tyr Gln Ala Gly Ser Asn Gly Gln
210 215 220

gtt gtc gtt gaa aag ggt ggc gag tgg cta ata aaa aat aat gat tcc 720
Val Val Val Glu Lys Gly Gly Glu Trp Leu Ile Lys Asn Asn Asp Ser
225 230 235 240

tca att gaa ttt caa att ggt aat caa gga act ggg gag gcg act att 768
Ser Ile Glu Phe Gln Ile Gly Asn Gln Gly Thr Gly Glu Ala Thr Ile
245 250 255

cgc gag ggg ggg ttg gtt acg gct gaa aat acg att atc ggt ggc aat 816
Arg Glu Gly Gly Leu Val Thr Ala Glu Asn Thr Ile Ile Gly Gly Asn
260 265 270

gcc acc ggt atc gga acc ctg aat gtg cag gat caa gac tct gtc atc 864
Ala Thr Gly Ile Gly Thr Leu Asn Val Gln Asp Gln Asp Ser Val Ile
275 280 285

acg gta cgc aga ctc tat aat gga tat ttc ggt aat ggc aca gtc aat 912
Thr Val Arg Arg Leu Tyr Asn Gly Tyr Phe Gly Asn Gly Thr Val Asn
290 295 300

att tcc aat aat gga ctg att aat aac aaa gaa tat tca ttg gtg ggc 960
Ile Ser Asn Asn Gly Leu Ile Asn Asn Lys Glu Tyr Ser Leu Val Gly
305 310 315 320

```
gtt cag gac ggt tcc cac ggt gtc gtc aac gtg acc gat aaa ggg cat    1008
Val Gln Asp Gly Ser His Gly Val Val Asn Val Thr Asp Lys Gly His
                325                 330                 335

tgg aat ttc ctc gga acg ggc gaa gct ttc cgc tat atc tat atc ggt    1056
Trp Asn Phe Leu Gly Thr Gly Glu Ala Phe Arg Tyr Ile Tyr Ile Gly
            340                 345                 350

gat gct ggc gac ggt gaa ctt aat gtc tcg agt gaa ggc aaa gta gat    1104
Asp Ala Gly Asp Gly Glu Leu Asn Val Ser Ser Glu Gly Lys Val Asp
        355                 360                 365

tcg gga att atc act gcg ggg atg aaa gaa aca ggc aca ggc aac att    1152
Ser Gly Ile Ile Thr Ala Gly Met Lys Glu Thr Gly Thr Gly Asn Ile
    370                 375                 380

act gtt aag gat aag aac tcc gtt atc act aat ctc gga act aat ctt    1200
Thr Val Lys Asp Lys Asn Ser Val Ile Thr Asn Leu Gly Thr Asn Leu
385                 390                 395                 400

ggt tat gac ggc cac ggc gaa atg aat atc agt aat cag ggg ctt gtt    1248
Gly Tyr Asp Gly His Gly Glu Met Asn Ile Ser Asn Gln Gly Leu Val
                405                 410                 415

gtc agc aac gga gga agt tca ctc ggt tat gga gaa acc ggc gtc ggg    1296
Val Ser Asn Gly Gly Ser Ser Leu Gly Tyr Gly Glu Thr Gly Val Gly
            420                 425                 430

aat gtc agc atc acc acg ggg gga atg tgg gag gtc aat aag aat gtc    1344
Asn Val Ser Ile Thr Thr Gly Gly Met Trp Glu Val Asn Lys Asn Val
        435                 440                 445

tat acc acc att ggt gtt gcg ggc gtc gga aac ctc aat att agc gat    1392
Tyr Thr Thr Ile Gly Val Ala Gly Val Gly Asn Leu Asn Ile Ser Asp
    450                 455                 460

ggc ggt aag ttc gta tcg caa aat att act ttt ttg ggc gat aaa gca    1440
Gly Gly Lys Phe Val Ser Gln Asn Ile Thr Phe Leu Gly Asp Lys Ala
465                 470                 475                 480

agc ggt atc ggc aca ctg aac ctg atg gat gcg aca tca tcg ttc gat    1488
Ser Gly Ile Gly Thr Leu Asn Leu Met Asp Ala Thr Ser Ser Phe Asp
                485                 490                 495

act gtg ggt atc aat gtc ggt aat ttt ggt agc ggt atc gta aat gtc    1536
Thr Val Gly Ile Asn Val Gly Asn Phe Gly Ser Gly Ile Val Asn Val
            500                 505                 510

agt aat ggt gcc acc ctt aat tca acg ggc tat gga ttt atc gga gga    1584
Ser Asn Gly Ala Thr Leu Asn Ser Thr Gly Tyr Gly Phe Ile Gly Gly
        515                 520                 525

aat gcc tcc ggt aag ggg ata gtt aat att tca acg gac agt ctc tgg    1632
Asn Ala Ser Gly Lys Gly Ile Val Asn Ile Ser Thr Asp Ser Leu Trp
    530                 535                 540

aat tta aag aca tca tct act aac gca caa ttg cta cag gtc ggt gta    1680
Asn Leu Lys Thr Ser Ser Thr Asn Ala Gln Leu Leu Gln Val Gly Val
545                 550                 555                 560

tta ggc acg ggt gaa ctg aat att acc acc gga ggt ata gtt aaa gcg    1728
```

```
Leu Gly Thr Gly Glu Leu Asn Ile Thr Thr Gly Gly Ile Val Lys Ala
                565                 570                 575

cgt gat aca cag ata gct ctc aat gac aaa agt aag ggc gac gtg agg    1776
Arg Asp Thr Gln Ile Ala Leu Asn Asp Lys Ser Lys Gly Asp Val Arg
            580                 585                 590

gtg gat ggg cag aac tct ctt ctt gaa aca ttc aat atg tac gta ggg    1824
Val Asp Gly Gln Asn Ser Leu Leu Glu Thr Phe Asn Met Tyr Val Gly
        595                 600                 605

aca tct ggt acg ggt acg tta acc ctg acg aat aac ggt acg ctg aat    1872
Thr Ser Gly Thr Gly Thr Leu Thr Leu Thr Asn Asn Gly Thr Leu Asn
    610                 615                 620

gtc gaa ggt gga gaa gtt tac tta ggt gtt ttt gag cct gct gta gga    1920
Val Glu Gly Gly Glu Val Tyr Leu Gly Val Phe Glu Pro Ala Val Gly
625                 630                 635                 640

acg cta aac att ggt gct gct cac ggt gag gcg gcg gca gat gcc gga    1968
Thr Leu Asn Ile Gly Ala Ala His Gly Glu Ala Ala Ala Asp Ala Gly
                645                 650                 655

ttt atc acc aat gcg acg aaa gtg gag ttt ggt ctt ggc gaa ggc gtt    2016
Phe Ile Thr Asn Ala Thr Lys Val Glu Phe Gly Leu Gly Glu Gly Val
            660                 665                 670

ttt gtc ttt aat cat acc aat aac agt gat gcc ggc tac cag gtc gat    2064
Phe Val Phe Asn His Thr Asn Asn Ser Asp Ala Gly Tyr Gln Val Asp
        675                 680                 685

atg ctg att aca ggt gac gat aaa gac gga aaa gtg atc cat gat gca    2112
Met Leu Ile Thr Gly Asp Asp Lys Asp Gly Lys Val Ile His Asp Ala
    690                 695                 700

ggc cat acg gtg ttc aat gca ggg aat act tat agc ggt aaa acg ctg    2160
Gly His Thr Val Phe Asn Ala Gly Asn Thr Tyr Ser Gly Lys Thr Leu
705                 710                 715                 720

gtc aat gac ggc ctc ctg acc att gcg tct cat acg gca gat ggg gta    2208
Val Asn Asp Gly Leu Leu Thr Ile Ala Ser His Thr Ala Asp Gly Val
                725                 730                 735

acg ggc atg ggg tcg agt gaa gta acc att gca aac ccc ggt acg ctc    2256
Thr Gly Met Gly Ser Ser Glu Val Thr Ile Ala Asn Pro Gly Thr Leu
            740                 745                 750

gac att ctc gca tca acg aac agt gca gga gat tac acg ctg acc aat    2304
Asp Ile Leu Ala Ser Thr Asn Ser Ala Gly Asp Tyr Thr Leu Thr Asn
        755                 760                 765

gcg ctc aaa ggc gat ggc ttg atg cga gtg cag ctg tca tcc tcc gac    2352
Ala Leu Lys Gly Asp Gly Leu Met Arg Val Gln Leu Ser Ser Ser Asp
    770                 775                 780

aag atg ttt ggc ttt acc cat gca aca ggg act gaa ttc gcc ggt gtt    2400
Lys Met Phe Gly Phe Thr His Ala Thr Gly Thr Glu Phe Ala Gly Val
785                 790                 795                 800

gcc caa ctg aaa gac agt acc ttc act ctg gaa cgc gac aac acc gct    2448
Ala Gln Leu Lys Asp Ser Thr Phe Thr Leu Glu Arg Asp Asn Thr Ala
```

805 810 815

gcg ctt act cac gcg atg ttg cag tct gac agt gaa aat acc aca tcg 2496
Ala Leu Thr His Ala Met Leu Gln Ser Asp Ser Glu Asn Thr Thr Ser
820 825 830

gta aaa gtt gga gag caa tcc att ggt gga ctg gcc atg aat gga ggt 2544
Val Lys Val Gly Glu Gln Ser Ile Gly Gly Leu Ala Met Asn Gly Gly
835 840 845

acc atc att ttc gat acg gat att cct gct gcg acg ctc gcg gag gga 2592
Thr Ile Ile Phe Asp Thr Asp Ile Pro Ala Ala Thr Leu Ala Glu Gly
850 855 860

tat atc agc gtc gat acg ctg gtt gtc ggc gcg ggt gac tac acc tgg 2640
Tyr Ile Ser Val Asp Thr Leu Val Val Gly Ala Gly Asp Tyr Thr Trp
865 870 875 880

aaa ggc cgt aac tat cag gta aac ggg acg ggc gac gtg ctt atc gac 2688
Lys Gly Arg Asn Tyr Gln Val Asn Gly Thr Gly Asp Val Leu Ile Asp
885 890 895

gtg cct aaa ccg tgg aat gat ccc atg gcg aat aac cct ctg acg acg 2736
Val Pro Lys Pro Trp Asn Asp Pro Met Ala Asn Asn Pro Leu Thr Thr
900 905 910

ctc aat ttg ctg gaa cac gac gat agc cat gtc ggc gtt caa ctg gtg 2784
Leu Asn Leu Leu Glu His Asp Asp Ser His Val Gly Val Gln Leu Val
915 920 925

aag gcg caa acg gtt att ggg tcg ggt ggc tca tta acg tta cgt gat 2832
Lys Ala Gln Thr Val Ile Gly Ser Gly Gly Ser Leu Thr Leu Arg Asp
930 935 940

tta cag ggc gac gag gtg gaa gcg gac aaa acg tta cac att gcg caa 2880
Leu Gln Gly Asp Glu Val Glu Ala Asp Lys Thr Leu His Ile Ala Gln
945 950 955 960

aac gga acg gtg gtc gcc gag ggt gat tat gga ttc cgc ctc acg acc 2928
Asn Gly Thr Val Val Ala Glu Gly Asp Tyr Gly Phe Arg Leu Thr Thr
965 970 975

gca cca ggt aat ggt ttg tac gtt aac tat ggg ctg aaa gcg ctg aac 2976
Ala Pro Gly Asn Gly Leu Tyr Val Asn Tyr Gly Leu Lys Ala Leu Asn
980 985 990

atc cat ggt ggg caa aag ctg acg tta gcc gaa cat ggc gga gcc tat 3024
Ile His Gly Gly Gln Lys Leu Thr Leu Ala Glu His Gly Gly Ala Tyr
995 1000 1005

ggc gca acg gcc gat atg tcg gca aaa atc ggt ggt gaa ggg gat ctg 3072
Gly Ala Thr Ala Asp Met Ser Ala Lys Ile Gly Gly Glu Gly Asp Leu
1010 1015 1020

gca atc aat acg gtg cga cag gtt tcg ctt tcc aac ggt cag aac gac 3120
Ala Ile Asn Thr Val Arg Gln Val Ser Leu Ser Asn Gly Gln Asn Asp
1025 1030 1035 1040

tat cag ggg gca acc tac gtt cag atg ggg aca tta cgt acc gat gcg 3168
Tyr Gln Gly Ala Thr Tyr Val Gln Met Gly Thr Leu Arg Thr Asp Ala
1045 1050 1055

gat ggt gca ctt ggc aac acc cgg gaa ctg aac atc agc aac gcg gcc 3216
Asp Gly Ala Leu Gly Asn Thr Arg Glu Leu Asn Ile Ser Asn Ala Ala
1060 1065 1070

atc gtc gat ctt aat gga tcg acg cag acg gta gag aca ttc acc ggg 3264
Ile Val Asp Leu Asn Gly Ser Thr Gln Thr Val Glu Thr Phe Thr Gly
1075 1080 1085

cag atg ggt tcg act gtt ttg ttc aaa gag ggg gcg ctg acg gta aat 3312
Gln Met Gly Ser Thr Val Leu Phe Lys Glu Gly Ala Leu Thr Val Asn
1090 1095 1100

aaa ggt ggg atc agt cag ggt gaa ctg aca ggt ggc gga aac ctg aat 3360
Lys Gly Gly Ile Ser Gln Gly Glu Leu Thr Gly Gly Gly Asn Leu Asn
1105 1110 1115 1120

gtt aca ggg gga acg ctg gct atc gag ggg ctt aat gca cgc tac aat 3408
Val Thr Gly Gly Thr Leu Ala Ile Glu Gly Leu Asn Ala Arg Tyr Asn
1125 1130 1135

gcg tta acc agc att agc cca aat gcg gaa gtc agc ctc gat aat act 3456
Ala Leu Thr Ser Ile Ser Pro Asn Ala Glu Val Ser Leu Asp Asn Thr
1140 1145 1150

cag ggg tta ggc aga gga aat att gcc aat gac ggt ctg tta acg cta 3504
Gln Gly Leu Gly Arg Gly Asn Ile Ala Asn Asp Gly Leu Leu Thr Leu
1155 1160 1165

aaa aac gtg act ggc gaa ctg cgt aat agc ata agc ggg aag ggt atc 3552
Lys Asn Val Thr Gly Glu Leu Arg Asn Ser Ile Ser Gly Lys Gly Ile
1170 1175 1180

gtg agc gca acc gcc agg aca gat gta gag ttg gat ggc gat aat agc 3600
Val Ser Ala Thr Ala Arg Thr Asp Val Glu Leu Asp Gly Asp Asn Ser
1185 1190 1195 1200

cgc ttt gtg ggg caa ttc aac att gat aca ggc agc gcg ctc agc gtc 3648
Arg Phe Val Gly Gln Phe Asn Ile Asp Thr Gly Ser Ala Leu Ser Val
1205 1210 1215

aac gag cag aaa aac ctg ggt gat gct tcc gtt atc aat aat ggc ctg 3696
Asn Glu Gln Lys Asn Leu Gly Asp Ala Ser Val Ile Asn Asn Gly Leu
1220 1225 1230

ctc acc atc tcc act gag cgt agc tgg gcg atg acg cac agt atc agc 3744
Leu Thr Ile Ser Thr Glu Arg Ser Trp Ala Met Thr His Ser Ile Ser
1235 1240 1245

ggt agc ggt gat gtg aca aaa ctg ggt acc ggg atc ctg act ctt aac 3792
Gly Ser Gly Asp Val Thr Lys Leu Gly Thr Gly Ile Leu Thr Leu Asn
1250 1255 1260

aac gat tcc gcg gcg tat cag ggt acg acg gat atc gtg ggg ggg gaa 3840
Asn Asp Ser Ala Ala Tyr Gln Gly Thr Thr Asp Ile Val Gly Gly Glu
1265 1270 1275 1280

att gct ttc ggt tcc gac tct gcc att aat atg gca agt caa cac att 3888
Ile Ala Phe Gly Ser Asp Ser Ala Ile Asn Met Ala Ser Gln His Ile
1285 1290 1295

aat atc cat aac agc ggt gtg atg tcg gga aat gtc acc act gca ggt 3936
Asn Ile His Asn Ser Gly Val Met Ser Gly Asn Val Thr Thr Ala Gly
1300 1305 1310

gat atg aac gtt atg cct ggg ggg ggc act gcg tgt cgc taa 3978
Asp Met Asn Val Met Pro Gly Gly Gly Thr Ala Cys Arg *
1315 1320 1325

<210> 112
<211> 756
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(756)

<400> 112
atg act gaa gca caa aga cat caa atc ctc ctg gaa atg ctc gca caa 48
Met Thr Glu Ala Gln Arg His Gln Ile Leu Leu Glu Met Leu Ala Gln
1 5 10 15

ttg ggc ttt gtg acc gtt gag aaa gtc gtt gag cgt ctg gga att tcg 96
Leu Gly Phe Val Thr Val Glu Lys Val Val Glu Arg Leu Gly Ile Ser
20 25 30

cct gcc act gcg cga cgc gat atc aat aaa ctt gac gaa agc ggc aaa 144
Pro Ala Thr Ala Arg Arg Asp Ile Asn Lys Leu Asp Glu Ser Gly Lys
35 40 45

ctg aaa aaa gtg cgc aat ggc gca gaa gct att acc caa cag cgc ccg 192
Leu Lys Lys Val Arg Asn Gly Ala Glu Ala Ile Thr Gln Gln Arg Pro
50 55 60

cgc tgg acg ccg atg aat ctg cat cag gcg cag aat cac gat gaa aaa 240
Arg Trp Thr Pro Met Asn Leu His Gln Ala Gln Asn His Asp Glu Lys
65 70 75 80

gta cgt atc gct aaa gcg gcc tcg cag ctg gtt aat ccg ggc gaa agc 288
Val Arg Ile Ala Lys Ala Ala Ser Gln Leu Val Asn Pro Gly Glu Ser
85 90 95

gta gtc atc aac tgc ggc tcc acc gcg ttt ctg ctt ggg cgg gaa atg 336
Val Val Ile Asn Cys Gly Ser Thr Ala Phe Leu Leu Gly Arg Glu Met
100 105 110

tgt ggc aag cca gtg caa atc atc act aat tat cta ccg ctg gca aat 384
Cys Gly Lys Pro Val Gln Ile Ile Thr Asn Tyr Leu Pro Leu Ala Asn
115 120 125

tac ctg atc gat caa gaa cat gac agc gtg atc att atg ggc gga cag 432
Tyr Leu Ile Asp Gln Glu His Asp Ser Val Ile Ile Met Gly Gly Gln
130 135 140

tac aac aaa agt cag tcc atc act tta agc ccg cag ggc agc gaa aac 480
Tyr Asn Lys Ser Gln Ser Ile Thr Leu Ser Pro Gln Gly Ser Glu Asn
145 150 155 160

agt ctc tat gcc ggg cac tgg atg ttt acc agc gga aaa ggg ctg acc 528

Ser Leu Tyr Ala Gly His Trp Met Phe Thr Ser Gly Lys Gly Leu Thr
165 170 175

gca gaa ggg ttg tat aaa acc gat atg ctg aca gca atg gca gag cag 576
Ala Glu Gly Leu Tyr Lys Thr Asp Met Leu Thr Ala Met Ala Glu Gln
180 185 190

aag atg ctg agc gtg gta ggg aaa ctg gtg gta ctg gtt gat agc agt 624
Lys Met Leu Ser Val Val Gly Lys Leu Val Val Leu Val Asp Ser Ser
195 200 205

aag att ggc gaa cgc gcg gga atg ctt ttt agc cgt gcc gat caa atc 672
Lys Ile Gly Glu Arg Ala Gly Met Leu Phe Ser Arg Ala Asp Gln Ile
210 215 220

gat atg ctt atc acc ggc aaa aat gct aac ccg gaa atc ctg caa caa 720
Asp Met Leu Ile Thr Gly Lys Asn Ala Asn Pro Glu Ile Leu Gln Gln
225 230 235 240

ctg gaa gcg caa ggg gtc agc att ctg cgt gtt taa 756
Leu Glu Ala Gln Gly Val Ser Ile Leu Arg Val *
245 250

<210> 113
<211> 825
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(825)

<400> 113
atg acc gaa ttt aca act ctt ctt cag caa gga aac gcc tgg ttc ttc 48
Met Thr Glu Phe Thr Thr Leu Leu Gln Gln Gly Asn Ala Trp Phe Phe
1 5 10 15

atc ccc agc gcc atc tta ctt ggt gcg ctt cat ggc ctg gaa cca ggg 96
Ile Pro Ser Ala Ile Leu Leu Gly Ala Leu His Gly Leu Glu Pro Gly
20 25 30

cac tca aaa acg atg atg gcg gcg ttt atc atc gcc atc aaa ggc acc 144
His Ser Lys Thr Met Met Ala Ala Phe Ile Ile Ala Ile Lys Gly Thr
35 40 45

att aaa caa gcg gtg atg ctc gga ctg gca gca act att tcg cat acc 192
Ile Lys Gln Ala Val Met Leu Gly Leu Ala Ala Thr Ile Ser His Thr
50 55 60

gca gtg gtc tgg tta att gcc ttt ggc ggg atg gtg atc agc aag cgc 240
Ala Val Val Trp Leu Ile Ala Phe Gly Gly Met Val Ile Ser Lys Arg
65 70 75 80

ttt act gct caa tca gca gaa ccg tgg ctc cag ctg att tcc gca gtg 288
Phe Thr Ala Gln Ser Ala Glu Pro Trp Leu Gln Leu Ile Ser Ala Val
85 90 95

atc att att agc acc gcg ttc tgg atg ttc tgg cgt acc tgg cgc ggc 336
Ile Ile Ile Ser Thr Ala Phe Trp Met Phe Trp Arg Thr Trp Arg Gly
100 105 110

```
gaa cgc aac tgg ctg gag aat atg cac ggg cat gat tat gag cat cat    384
Glu Arg Asn Trp Leu Glu Asn Met His Gly His Asp Tyr Glu His His
        115                 120                 125

cat cac gat cac gaa cat cac cac gac cat gga cat cat cac cat cac    432
His His Asp His Glu His His His Asp His Gly His His His His His
    130                 135                 140

gaa cat ggc gag tat cag gat gcc cat gca cga gcc cat gcc aat gac    480
Glu His Gly Glu Tyr Gln Asp Ala His Ala Arg Ala His Ala Asn Asp
145                 150                 155                 160

att aaa cga cgc ttt gat ggt aga gag gtc acc aac tgg caa att ttg    528
Ile Lys Arg Arg Phe Asp Gly Arg Glu Val Thr Asn Trp Gln Ile Leu
                165                 170                 175

tta ttt ggc tta acc ggt ggc ctt atc ccc tgc ccg gca gca att acc    576
Leu Phe Gly Leu Thr Gly Gly Leu Ile Pro Cys Pro Ala Ala Ile Thr
            180                 185                 190

gtg ctg ttg ata tgc att cag ttg aaa gcc ctg aca ctg ggc gca aca    624
Val Leu Leu Ile Cys Ile Gln Leu Lys Ala Leu Thr Leu Gly Ala Thr
        195                 200                 205

ctg gtc gtc agt ttc agc att ggc ctg gcg tta acg ctt gtc acc gta    672
Leu Val Val Ser Phe Ser Ile Gly Leu Ala Leu Thr Leu Val Thr Val
    210                 215                 220

ggc gtt ggc gca gca atc agc gtt cag cag gtc gca aaa cgc tgg agc    720
Gly Val Gly Ala Ala Ile Ser Val Gln Gln Val Ala Lys Arg Trp Ser
225                 230                 235                 240

gga ttt aac act ctc gct aaa cgc gcc ccc tat ttt tcc agt ctg ttg    768
Gly Phe Asn Thr Leu Ala Lys Arg Ala Pro Tyr Phe Ser Ser Leu Leu
                245                 250                 255

att ggc tta gtc ggt gtg tat atg ggc gta cat ggc ttc atg ggc ata    816
Ile Gly Leu Val Gly Val Tyr Met Gly Val His Gly Phe Met Gly Ile
            260                 265                 270

atg cga taa                                                        825
Met Arg  *
```

<210> 114
<211> 519
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(519)

<400> 114

```
atg att ctt aaa tca gct att tcc gct gac tct ctg ctc gct aag gat    48
Met Ile Leu Lys Ser Ala Ile Ser Ala Asp Ser Leu Leu Ala Lys Asp
 1               5                  10                  15

gct ttt agg gca tcc ttt cat tta cac ttt tta cga aat cat ggg atc    96
```

```
Ala Phe Arg Ala Ser Phe His Leu His Phe Leu Arg Asn His Gly Ile
            20                  25                  30

act aac aaa ata tcg ctt gtc agt tat att gta tgg cag gaa aga tat      144
Thr Asn Lys Ile Ser Leu Val Ser Tyr Ile Val Trp Gln Glu Arg Tyr
        35                  40                  45

gcg act gat att aca gat ccc caa agt gga gag ttt atg acc att aaa      192
Ala Thr Asp Ile Thr Asp Pro Gln Ser Gly Glu Phe Met Thr Ile Lys
    50                  55                  60

aat aag atg ttg ctg ggt gcg ctt ttg ctg gtt acc agt gcc gcc tgg      240
Asn Lys Met Leu Leu Gly Ala Leu Leu Leu Val Thr Ser Ala Ala Trp
65                  70                  75                  80

gcc gca cca gcc acc gcg ggt tcg acc aat acc tcg gga att tct aag      288
Ala Ala Pro Ala Thr Ala Gly Ser Thr Asn Thr Ser Gly Ile Ser Lys
                85                  90                  95

tat gag tta agt agt ttc att gct gac ttt aag cat ttc aaa cca ggg      336
Tyr Glu Leu Ser Ser Phe Ile Ala Asp Phe Lys His Phe Lys Pro Gly
            100                 105                 110

gac acc gta cca gaa atg tac cgt acc gat gag tac aac att aag cag      384
Asp Thr Val Pro Glu Met Tyr Arg Thr Asp Glu Tyr Asn Ile Lys Gln
        115                 120                 125

tgg cag ttg cgt aac ctg ccc gcg cct gat gcc ggg acg cac tgg acc      432
Trp Gln Leu Arg Asn Leu Pro Ala Pro Asp Ala Gly Thr His Trp Thr
    130                 135                 140

tat atg ggt ggc gcg tac gtg ttg atc agc gac acc gac ggt aaa atc      480
Tyr Met Gly Gly Ala Tyr Val Leu Ile Ser Asp Thr Asp Gly Lys Ile
145                 150                 155                 160

att aaa gcc tac gac ggt gag att ttt tat cat cgc taa                  519
Ile Lys Ala Tyr Asp Gly Glu Ile Phe Tyr His Arg  *
                165                 170


&lt;210&gt; 115
&lt;211&gt; 1035
&lt;212&gt; DNA
&lt;213&gt; Escherichia coli

&lt;220&gt;
&lt;221&gt; CDS
&lt;222&gt; (1)...(1035)

&lt;400&gt; 115
atg gaa att cgc ata atg cta ttt ata tta atg atg atg gtt atg cct      48
Met Glu Ile Arg Ile Met Leu Phe Ile Leu Met Met Met Val Met Pro
 1               5                  10                  15

gtg agc tat gcg gca tgt tat agt gag tta tct gtt cag cac aac ttg      96
Val Ser Tyr Ala Ala Cys Tyr Ser Glu Leu Ser Val Gln His Asn Leu
            20                  25                  30

gtt gtt cag ggg gat ttt gca ctt act caa aca caa atg gcg aca tat      144
Val Val Gln Gly Asp Phe Ala Leu Thr Gln Thr Gln Met Ala Thr Tyr
        35                  40                  45
```

gag cat aat ttt aat gat tcg tca tgc gta agt aca aat act atc acc 192
Glu His Asn Phe Asn Asp Ser Ser Cys Val Ser Thr Asn Thr Ile Thr
50 55 60

cct atg agc ccg tcg gat att att gtt gga ctt tat aac gat acc ata 240
Pro Met Ser Pro Ser Asp Ile Ile Val Gly Leu Tyr Asn Asp Thr Ile
65 70 75 80

aaa tta aat tta cat ttt gaa tgg acc aat aaa aac aac atc acg ttg 288
Lys Leu Asn Leu His Phe Glu Trp Thr Asn Lys Asn Asn Ile Thr Leu
85 90 95

tca aat aat cag acc agt ttc acc agt ggt tat tca gtt acg gtg aca 336
Ser Asn Asn Gln Thr Ser Phe Thr Ser Gly Tyr Ser Val Thr Val Thr
100 105 110

cct gcg gcc agt aat gca aaa gtg aat gtt tct gcg ggg ggc ggc ggt 384
Pro Ala Ala Ser Asn Ala Lys Val Asn Val Ser Ala Gly Gly Gly Gly
115 120 125

tca gtg atg att aat ggt gtt gcg aca tta tcc agt gct tca tca tcg 432
Ser Val Met Ile Asn Gly Val Ala Thr Leu Ser Ser Ala Ser Ser Ser
130 135 140

aca cgc ggg agt gcc gca gta caa ttt cta ctg tgt tta tta ggt ggc 480
Thr Arg Gly Ser Ala Ala Val Gln Phe Leu Leu Cys Leu Leu Gly Gly
145 150 155 160

aag tca tgg gat gca tgt gta aat agc tac aga aat gca ttg gca caa 528
Lys Ser Trp Asp Ala Cys Val Asn Ser Tyr Arg Asn Ala Leu Ala Gln
165 170 175

aat gca ggt gtc tat tcc ttt aat ctg aca ttg tca tac aac ccg ata 576
Asn Ala Gly Val Tyr Ser Phe Asn Leu Thr Leu Ser Tyr Asn Pro Ile
180 185 190

acc aca acc tgc aaa ccg gac gat tta tta att act tta gac agt att 624
Thr Thr Thr Cys Lys Pro Asp Asp Leu Leu Ile Thr Leu Asp Ser Ile
195 200 205

ccc gtt tca caa tta cca gcc aca ggt aac aaa gca aca ata aat agt 672
Pro Val Ser Gln Leu Pro Ala Thr Gly Asn Lys Ala Thr Ile Asn Ser
210 215 220

aaa caa ggg gat att att ctg cgt tgt aaa aat tta tta ggt caa caa 720
Lys Gln Gly Asp Ile Ile Leu Arg Cys Lys Asn Leu Leu Gly Gln Gln
225 230 235 240

aat caa aca tca cgg aaa atg cag gtg tat tta tca agt tct gac ttg 768
Asn Gln Thr Ser Arg Lys Met Gln Val Tyr Leu Ser Ser Ser Asp Leu
245 250 255

tta acc aac agc aac aca ata ctg aaa ggt gcg gaa gat aat ggc gta 816
Leu Thr Asn Ser Asn Thr Ile Leu Lys Gly Ala Glu Asp Asn Gly Val
260 265 270

gga ttt att ctt gaa agt aat ggt tcg cca gtc aca ctt tta aat atc 864
Gly Phe Ile Leu Glu Ser Asn Gly Ser Pro Val Thr Leu Leu Asn Ile
275 280 285

act aac agc agt aaa gga tat aca aat tta aag gaa gtt gcg gcg aag 912
Thr Asn Ser Ser Lys Gly Tyr Thr Asn Leu Lys Glu Val Ala Ala Lys
290 295 300

tca aaa ctt aca gat aca acg gtt tca att ccg ata aca gcc agt tac 960
Ser Lys Leu Thr Asp Thr Thr Val Ser Ile Pro Ile Thr Ala Ser Tyr
305 310 315 320

tac gtc tac gat aca aac aaa gtt aaa tct ggc gca ctg gag gca acc 1008
Tyr Val Tyr Asp Thr Asn Lys Val Lys Ser Gly Ala Leu Glu Ala Thr
325 330 335

gca tta atc aac gtg aaa tac gac taa 1035
Ala Leu Ile Asn Val Lys Tyr Asp *
340

<210> 116
<211> 2481
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(2481)

<400> 116
atg ttg aga atg acc cca ctt gca tca gca atc gta gcg tta ttg ctc 48
Met Leu Arg Met Thr Pro Leu Ala Ser Ala Ile Val Ala Leu Leu Leu
1 5 10 15

ggc att gaa gct tat gca gct gaa gaa acc ttt gat acc cat ttt atg 96
Gly Ile Glu Ala Tyr Ala Ala Glu Glu Thr Phe Asp Thr His Phe Met
20 25 30

ata ggt gga atg aaa gac cag cag gtt gca aat att cgt ctt gat gat 144
Ile Gly Gly Met Lys Asp Gln Gln Val Ala Asn Ile Arg Leu Asp Asp
35 40 45

aat caa ccc tta ccg ggg cag tat gac atc gat att tat gtc aat aag 192
Asn Gln Pro Leu Pro Gly Gln Tyr Asp Ile Asp Ile Tyr Val Asn Lys
50 55 60

caa tgg cgc ggg aaa tat gag att att gtt aaa gac aac ccg caa gaa 240
Gln Trp Arg Gly Lys Tyr Glu Ile Ile Val Lys Asp Asn Pro Gln Glu
65 70 75 80

aca tgt tta tca aga gaa gtt atc aag cgg tta ggc att aat agc gat 288
Thr Cys Leu Ser Arg Glu Val Ile Lys Arg Leu Gly Ile Asn Ser Asp
85 90 95

aac ttc gcc agc ggt aag caa tgt tta aca ttt gag caa ctt gtt cag 336
Asn Phe Ala Ser Gly Lys Gln Cys Leu Thr Phe Glu Gln Leu Val Gln
100 105 110

ggt ggg agc tat acc tgg gat atc ggg gtt ttt cgt ctc gat ttc agt 384
Gly Gly Ser Tyr Thr Trp Asp Ile Gly Val Phe Arg Leu Asp Phe Ser
115 120 125

gtc ccg cag gcc tgg gtg gaa gaa ctg gaa agt ggc tat gtt cca ccg 432
Val Pro Gln Ala Trp Val Glu Glu Leu Glu Ser Gly Tyr Val Pro Pro

```
    130                 135                 140

gaa aac tgg gag cgg ggt att aat gcg ttt tat acc tct tat tat ctg     480
Glu Asn Trp Glu Arg Gly Ile Asn Ala Phe Tyr Thr Ser Tyr Tyr Leu
145                 150                 155                 160

agt cag tat tac agc gac tat aaa gcg tcg ggt aat aac aag agt aca     528
Ser Gln Tyr Tyr Ser Asp Tyr Lys Ala Ser Gly Asn Asn Lys Ser Thr
                165                 170                 175

tat gta cgt ttt aac agc ggg tta aat tta ctg ggg tgg caa ctg cat     576
Tyr Val Arg Phe Asn Ser Gly Leu Asn Leu Leu Gly Trp Gln Leu His
            180                 185                 190

tct gat gcc agt ttc agt aaa aca aat aac aat cca ggg gtg tgg aaa     624
Ser Asp Ala Ser Phe Ser Lys Thr Asn Asn Asn Pro Gly Val Trp Lys
        195                 200                 205

agc aat acc ctg tat ctg gaa cgt gga ttt gcc caa ctt ctc ggc acg     672
Ser Asn Thr Leu Tyr Leu Glu Arg Gly Phe Ala Gln Leu Leu Gly Thr
    210                 215                 220

ctt cgc gtg ggt gat atg tac aca tca agc gat att ttt gat tct gtt     720
Leu Arg Val Gly Asp Met Tyr Thr Ser Ser Asp Ile Phe Asp Ser Val
225                 230                 235                 240

cgc ttc aga ggt gtg cgg ttg ttt cgt gat atg cag atg ttg cct aac     768
Arg Phe Arg Gly Val Arg Leu Phe Arg Asp Met Gln Met Leu Pro Asn
                245                 250                 255

tcg aaa caa aat ttt acg cca cgg gtg cag ggg att gct cag agt aac     816
Ser Lys Gln Asn Phe Thr Pro Arg Val Gln Gly Ile Ala Gln Ser Asn
            260                 265                 270

gcg ctg gta act att gaa cag aat ggt ttt gtg gtt tat cag aaa gag     864
Ala Leu Val Thr Ile Glu Gln Asn Gly Phe Val Val Tyr Gln Lys Glu
        275                 280                 285

gtt cct cct ggc ccg ttc gcg att aca gat ttg cag ttg gcc ggt ggt     912
Val Pro Pro Gly Pro Phe Ala Ile Thr Asp Leu Gln Leu Ala Gly Gly
    290                 295                 300

gga gca gat ctt gat gtc agc gtg aaa gag gcg gac ggc tcg gta acc     960
Gly Ala Asp Leu Asp Val Ser Val Lys Glu Ala Asp Gly Ser Val Thr
305                 310                 315                 320

acc tat ctg gtg cct tat gca gcg gtg cca aat atg ctg caa ccc ggc    1008
Thr Tyr Leu Val Pro Tyr Ala Ala Val Pro Asn Met Leu Gln Pro Gly
                325                 330                 335

gtg tcg aaa tat gat tta gcg gcg ggt cgt agc cat att gaa ggg gcg    1056
Val Ser Lys Tyr Asp Leu Ala Ala Gly Arg Ser His Ile Glu Gly Ala
            340                 345                 350

agc aaa caa agt gat ttt gtc cag gcg ggt tat cag tat ggt ttt aat    1104
Ser Lys Gln Ser Asp Phe Val Gln Ala Gly Tyr Gln Tyr Gly Phe Asn
        355                 360                 365

aat tta ttg acg ctg tat ggt ggc tcg atg gtc gcg aat aat tat tac    1152
Asn Leu Leu Thr Leu Tyr Gly Gly Ser Met Val Ala Asn Asn Tyr Tyr
    370                 375                 380
```

gcg ttt act ttg ggg gct ggc tgg aat aca cgc att ggt gcc att tcc 1200
Ala Phe Thr Leu Gly Ala Gly Trp Asn Thr Arg Ile Gly Ala Ile Ser
385 390 395 400

gtc gat gcc act aag tcg cat agt aaa caa gac aac ggc gat gtg ttt 1248
Val Asp Ala Thr Lys Ser His Ser Lys Gln Asp Asn Gly Asp Val Phe
405 410 415

gac ggg caa agt tat caa att gcc tac aac aaa ttt gtg agc caa acg 1296
Asp Gly Gln Ser Tyr Gln Ile Ala Tyr Asn Lys Phe Val Ser Gln Thr
420 425 430

tcg acg cgt ttt ggt ctg gcg gcc tgg cgt tat tcg tcg cgt gat tac 1344
Ser Thr Arg Phe Gly Leu Ala Ala Trp Arg Tyr Ser Ser Arg Asp Tyr
435 440 445

cgg aca ttt aac gat cac gtt tgg gca aac aat aaa gat aat tat cgc 1392
Arg Thr Phe Asn Asp His Val Trp Ala Asn Asn Lys Asp Asn Tyr Arg
450 455 460

cgt gat gaa aac gat gtc tat gac att gcc gat tat tac cag aac gat 1440
Arg Asp Glu Asn Asp Val Tyr Asp Ile Ala Asp Tyr Tyr Gln Asn Asp
465 470 475 480

ttt ggc cgc aaa aat agc ttt tcc gcc aat atg agc cag tca ttg cca 1488
Phe Gly Arg Lys Asn Ser Phe Ser Ala Asn Met Ser Gln Ser Leu Pro
485 490 495

gaa ggt tgg ggg tct gtg tca tta agt acg tta tgg cga gat tac tgg 1536
Glu Gly Trp Gly Ser Val Ser Leu Ser Thr Leu Trp Arg Asp Tyr Trp
500 505 510

ggg cgt agc ggc agt agt aag gat tat cag ttg agt tat tcc aac aac 1584
Gly Arg Ser Gly Ser Ser Lys Asp Tyr Gln Leu Ser Tyr Ser Asn Asn
515 520 525

ctg cga cgg ata agc tat acc ctc gcg gca agc cag gct tat gac gag 1632
Leu Arg Arg Ile Ser Tyr Thr Leu Ala Ala Ser Gln Ala Tyr Asp Glu
530 535 540

aat cat cat gaa gag aaa cgt ttt aat att ttt ata tcg att ccc ttt 1680
Asn His His Glu Glu Lys Arg Phe Asn Ile Phe Ile Ser Ile Pro Phe
545 550 555 560

gat tgg ggt gat gac gtt tcg acg cct cgt cgg caa ata tat atg tct 1728
Asp Trp Gly Asp Asp Val Ser Thr Pro Arg Arg Gln Ile Tyr Met Ser
565 570 575

aac tca acg acg ttt gat gat cag ggg ttt gcc tca aat aat acg gga 1776
Asn Ser Thr Thr Phe Asp Asp Gln Gly Phe Ala Ser Asn Asn Thr Gly
580 585 590

tta tca gga aca gta ggg agt cgg gat cag ttc aat tat ggt gtc aac 1824
Leu Ser Gly Thr Val Gly Ser Arg Asp Gln Phe Asn Tyr Gly Val Asn
595 600 605

ctg agt cat caa cat cag gga aat gaa acg aca gct ggg gcg aat ttg 1872
Leu Ser His Gln His Gln Gly Asn Glu Thr Thr Ala Gly Ala Asn Leu
610 615 620

```
acc tgg aac gcg ccg gtt gcg aca gtg aat ggc agt tat agt cag tcg    1920
Thr Trp Asn Ala Pro Val Ala Thr Val Asn Gly Ser Tyr Ser Gln Ser
625                 630                 635                 640

agt act tat cga cag gct gga gcc agt gtt tca ggg ggc att gtc gcc    1968
Ser Thr Tyr Arg Gln Ala Gly Ala Ser Val Ser Gly Gly Ile Val Ala
                645                 650                 655

tgg tcg ggt ggc gtt aat ctg gcg aac cgt ctt tcc gaa acg ttt gct    2016
Trp Ser Gly Gly Val Asn Leu Ala Asn Arg Leu Ser Glu Thr Phe Ala
            660                 665                 670

gtg atg aat gcg cca gga att aaa gat gct tat gtc aat ggg caa aaa    2064
Val Met Asn Ala Pro Gly Ile Lys Asp Ala Tyr Val Asn Gly Gln Lys
        675                 680                 685

tat cgc aca aca aac cgt aat gga gtg gtg ata tac gac gga atg aca    2112
Tyr Arg Thr Thr Asn Arg Asn Gly Val Val Ile Tyr Asp Gly Met Thr
    690                 695                 700

cct tat cgg gaa aat cac ctg atg ctg gat gtg tcg caa agc gat agc    2160
Pro Tyr Arg Glu Asn His Leu Met Leu Asp Val Ser Gln Ser Asp Ser
705                 710                 715                 720

gaa gca gaa tta cgt ggc aac cgg aaa att gcc gcc cct tat cgc ggc    2208
Glu Ala Glu Leu Arg Gly Asn Arg Lys Ile Ala Ala Pro Tyr Arg Gly
                725                 730                 735

gcg gtt gta ctg gtt aat ttt gat acc gat cag cgc aag cca tgg ttt    2256
Ala Val Val Leu Val Asn Phe Asp Thr Asp Gln Arg Lys Pro Trp Phe
            740                 745                 750

ata aaa gcg tta aga gca gat ggg caa tca tta acg ttt ggt tat gaa    2304
Ile Lys Ala Leu Arg Ala Asp Gly Gln Ser Leu Thr Phe Gly Tyr Glu
        755                 760                 765

gtc aat gat atc cat ggt cat aat att ggc gtt gtc ggc cag gga agt    2352
Val Asn Asp Ile His Gly His Asn Ile Gly Val Val Gly Gln Gly Ser
    770                 775                 780

cag tta ttt att cgc acc aat gaa gta ccg cca tcg gtt aat gtg gca    2400
Gln Leu Phe Ile Arg Thr Asn Glu Val Pro Pro Ser Val Asn Val Ala
785                 790                 795                 800

att gat aag caa caa gga ctt tca tgc aca atc acc ttc ggt aaa gag    2448
Ile Asp Lys Gln Gln Gly Leu Ser Cys Thr Ile Thr Phe Gly Lys Glu
                805                 810                 815

att gat gaa agt aga aat tat att tgc cag taa                        2481
Ile Asp Glu Ser Arg Asn Tyr Ile Cys Gln  *
            820                 825
```

<210> 117
<211> 720
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(720)

```
<400> 117
atg gcc gct atc cca tgg cgg cct ttt aat tta aga ggc att aaa atg      48
Met Ala Ala Ile Pro Trp Arg Pro Phe Asn Leu Arg Gly Ile Lys Met
 1               5                  10                  15

aaa gga tta tta tct tta ctc att ttt tct atg gtc ctt cct gca cat      96
Lys Gly Leu Leu Ser Leu Leu Ile Phe Ser Met Val Leu Pro Ala His
            20                  25                  30

gcc gga att gtt atc tac ggg acg cgc att att tac ccg gca gaa aat     144
Ala Gly Ile Val Ile Tyr Gly Thr Arg Ile Ile Tyr Pro Ala Glu Asn
        35                  40                  45

aaa gaa gtg atg gtg cag ttg atg aac cag gga aac cgt tct tcg ctg     192
Lys Glu Val Met Val Gln Leu Met Asn Gln Gly Asn Arg Ser Ser Leu
    50                  55                  60

ctg cag gcg tgg att gat gat ggc gat acg tca tta cca cca gaa aaa     240
Leu Gln Ala Trp Ile Asp Asp Gly Asp Thr Ser Leu Pro Pro Glu Lys
 65                  70                  75                  80

att cag gtt cct ttc atg tta acg cca cca gtg gca aaa ata ggg gca     288
Ile Gln Val Pro Phe Met Leu Thr Pro Pro Val Ala Lys Ile Gly Ala
                85                  90                  95

aat tcc ggg cag caa gta aaa atc aaa att atg ccg aat aaa ctg ccc     336
Asn Ser Gly Gln Gln Val Lys Ile Lys Ile Met Pro Asn Lys Leu Pro
            100                 105                 110

act aat aaa gaa agc att ttt tat ctg aat gtt ctg gat att cca cca     384
Thr Asn Lys Glu Ser Ile Phe Tyr Leu Asn Val Leu Asp Ile Pro Pro
        115                 120                 125

aat agt cca gag caa gaa ggt aag aat gca ctg aag ttt gcg atg caa     432
Asn Ser Pro Glu Gln Glu Gly Lys Asn Ala Leu Lys Phe Ala Met Gln
    130                 135                 140

aac aga att aag ttg ttt tac cgg cca gcg ggt att gct ccg gta aat     480
Asn Arg Ile Lys Leu Phe Tyr Arg Pro Ala Gly Ile Ala Pro Val Asn
145                 150                 155                 160

aaa gcg aca ttt aaa aaa ttg ctg gta aat cgc agt ggc aat ggt ttg     528
Lys Ala Thr Phe Lys Lys Leu Leu Val Asn Arg Ser Gly Asn Gly Leu
                165                 170                 175

gtg ata aaa aat gac tca gct aat tgg gtg acg att tcg gat gtc aaa     576
Val Ile Lys Asn Asp Ser Ala Asn Trp Val Thr Ile Ser Asp Val Lys
            180                 185                 190

gct aat aat gtc aaa gtc aat tat gaa act att atg att gcc ccc tta     624
Ala Asn Asn Val Lys Val Asn Tyr Glu Thr Ile Met Ile Ala Pro Leu
        195                 200                 205

gaa agt cag agt gtt aat gtc aaa agt aat aat gca aat aac tgg cat     672
Glu Ser Gln Ser Val Asn Val Lys Ser Asn Asn Ala Asn Asn Trp His
    210                 215                 220

ctg acc att atc gat gac cat ggc aac tat att agt gac aaa att taa     720
Leu Thr Ile Ile Asp Asp His Gly Asn Tyr Ile Ser Asp Lys Ile  *
225                 230                 235
```

<210> 118
<211> 543
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(543)

<400> 118
```
atg aaa cgt tca att att gct gcc gct gtc ttt tct tct ttt ttt atg      48
Met Lys Arg Ser Ile Ile Ala Ala Ala Val Phe Ser Ser Phe Phe Met
 1               5                  10                  15

agc gct gga gta ttt gct gca gac gtt gat acc gga aca tta act att      96
Ser Ala Gly Val Phe Ala Ala Asp Val Asp Thr Gly Thr Leu Thr Ile
            20                  25                  30

aag ggg aat att gca gaa tct ccg tgt aaa ttc gaa gcg ggt ggt gat      144
Lys Gly Asn Ile Ala Glu Ser Pro Cys Lys Phe Glu Ala Gly Gly Asp
        35                  40                  45

tca gta agt att aat atg ccg act gta cca acc agt gtc ttt gaa ggt      192
Ser Val Ser Ile Asn Met Pro Thr Val Pro Thr Ser Val Phe Glu Gly
    50                  55                  60

aaa gct aaa tat tct acc tat gat gat gca gtc ggt gta acc agc agc      240
Lys Ala Lys Tyr Ser Thr Tyr Asp Asp Ala Val Gly Val Thr Ser Ser
 65                 70                  75                      80

atg tta aaa att agc tgc ccg aaa gaa gtt gct ggt gta aaa ctc tcg      288
Met Leu Lys Ile Ser Cys Pro Lys Glu Val Ala Gly Val Lys Leu Ser
                85                  90                  95

ttg att acc aac gat aaa ata acc ggt aac gat aag gcg ata gcc agt      336
Leu Ile Thr Asn Asp Lys Ile Thr Gly Asn Asp Lys Ala Ile Ala Ser
            100                 105                 110

agc aac gat acc gtg ggt tac tat ctc tat tta ggt gat aac agc gat      384
Ser Asn Asp Thr Val Gly Tyr Tyr Leu Tyr Leu Gly Asp Asn Ser Asp
        115                 120                 125

gtc ctg gat gtt tct gca cct ttt aac att gag agt tat aaa aca gcg      432
Val Leu Asp Val Ser Ala Pro Phe Asn Ile Glu Ser Tyr Lys Thr Ala
    130                 135                 140

gaa ggt caa tat gct att ccg ttt aaa gca aaa tac ctg aaa ctg aca      480
Glu Gly Gln Tyr Ala Ile Pro Phe Lys Ala Lys Tyr Leu Lys Leu Thr
145                 150                 155                 160

gat aac tca gtg caa tca ggt gat gtg tta tct tct ctg gtt atg cgt      528
Asp Asn Ser Val Gln Ser Gly Asp Val Leu Ser Ser Leu Val Met Arg
                165                 170                 175

gtg gcg cag gat taa                                                  543
Val Ala Gln Asp  *
            180
```

<210> 119
<211> 288
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(288)

<400> 119
atg ctg cac aca tta cat cgc tca ccc tgg ctg acg gat ttt gct gcg 48
Met Leu His Thr Leu His Arg Ser Pro Trp Leu Thr Asp Phe Ala Ala
1 5 10 15

ctg ctg cgt ctg ctc agt gaa gga gac gaa ctg cta tta ttg caa gat 96
Leu Leu Arg Leu Leu Ser Glu Gly Asp Glu Leu Leu Leu Leu Gln Asp
20 25 30

ggc gta act gcc gca gtt gac ggt aac cgc tac ctt gaa agt ctg cgt 144
Gly Val Thr Ala Ala Val Asp Gly Asn Arg Tyr Leu Glu Ser Leu Arg
35 40 45

aat gcc ccc att aag gtc tat gcc ctg aac gaa gac ctt att gcc cgc 192
Asn Ala Pro Ile Lys Val Tyr Ala Leu Asn Glu Asp Leu Ile Ala Arg
50 55 60

ggt ttg act ggt caa att tcg aac gac atc att ctc att gac tat act 240
Gly Leu Thr Gly Gln Ile Ser Asn Asp Ile Ile Leu Ile Asp Tyr Thr
65 70 75 80

gat ttc gtc aga ctt acg gtt aag cac ccc agc cag atg gcc tgg tga 288
Asp Phe Val Arg Leu Thr Val Lys His Pro Ser Gln Met Ala Trp *
85 90 95

<210> 120
<211> 360
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(360)

<400> 120
atg aaa cga att gcg ttt gtt ttt tct act gca cct cat ggt aca gcc 48
Met Lys Arg Ile Ala Phe Val Phe Ser Thr Ala Pro His Gly Thr Ala
1 5 10 15

gca ggc cgg gaa ggt tta gat gct tta ctg gca act tcc gca tta act 96
Ala Gly Arg Glu Gly Leu Asp Ala Leu Leu Ala Thr Ser Ala Leu Thr
20 25 30

gac gat ctg gct gtc ttc ttt ata gct gat ggc gtt ttt cag ctg ctg 144
Asp Asp Leu Ala Val Phe Phe Ile Ala Asp Gly Val Phe Gln Leu Leu
35 40 45

cca gga caa aag ccc gat gca gtg ctg gcg cgt gat tac att gcc act 192
Pro Gly Gln Lys Pro Asp Ala Val Leu Ala Arg Asp Tyr Ile Ala Thr

```
    50                  55                  60

ttt aaa ttg ttg ggt ctg tac gac att gaa cag tgc tgg gtt tgt gcg      240
Phe Lys Leu Leu Gly Leu Tyr Asp Ile Glu Gln Cys Trp Val Cys Ala
 65                  70                  75                  80

gct tca ctg cgc gaa cgc ggg tta gat ccg cag aca ccc ttt gtt gtc      288
Ala Ser Leu Arg Glu Arg Gly Leu Asp Pro Gln Thr Pro Phe Val Val
                 85                  90                  95

gaa gcc acg ccg ctc gaa gca gat gcc tta cgc cgc gaa ctc gcc aac      336
Glu Ala Thr Pro Leu Glu Ala Asp Ala Leu Arg Arg Glu Leu Ala Asn
            100                 105                 110

tac gat gtt att ttg agg ttt tga                                      360
Tyr Asp Val Ile Leu Arg Phe  *
        115


<210> 121
<211> 387
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(387)

<400> 121
atg cgt ttt gcc atc gtg gtg acc ggg cca gca tac ggt acg caa cag      48
Met Arg Phe Ala Ile Val Val Thr Gly Pro Ala Tyr Gly Thr Gln Gln
 1               5                  10                  15

gcg agt agt gct ttt cag ttt gcg cag gcg ctg ata gca gat ggc cat      96
Ala Ser Ser Ala Phe Gln Phe Ala Gln Ala Leu Ile Ala Asp Gly His
            20                  25                  30

gag tta agc agc gtc ttt ttc tat cgg gaa ggg gtc tat aac gct aac      144
Glu Leu Ser Ser Val Phe Phe Tyr Arg Glu Gly Val Tyr Asn Ala Asn
        35                  40                  45

caa ttg acc tct ccg gca agt gac gaa ttt gac ctc gta cgg gcc tgg      192
Gln Leu Thr Ser Pro Ala Ser Asp Glu Phe Asp Leu Val Arg Ala Trp
    50                  55                  60

caa caa ctg aat gcg caa cat ggt gtg gcg ctg aat atc tgc gta gcg      240
Gln Gln Leu Asn Ala Gln His Gly Val Ala Leu Asn Ile Cys Val Ala
 65                  70                  75                  80

gca gca tta cgc cgt ggc gtt gtt gat gaa acg gag gcc gga aga ctg      288
Ala Ala Leu Arg Arg Gly Val Val Asp Glu Thr Glu Ala Gly Arg Leu
                 85                  90                  95

ggg ctg gct tcg tca aac ctt cag cag gga ttt acc tta agc gga ctt      336
Gly Leu Ala Ser Ser Asn Leu Gln Gln Gly Phe Thr Leu Ser Gly Leu
            100                 105                 110

ggg gcg ctg gcg gaa gcc tcg ctg acc tgt gac agg gtg gta cag ttc      384
Gly Ala Leu Ala Glu Ala Ser Leu Thr Cys Asp Arg Val Val Gln Phe
        115                 120                 125
```

tga 387
*

<210> 122
<211> 735
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(735)

<400> 122
gtg ttt ttt ttc atg tcc agg tcg ctt tta acc aac gaa acc agt gag 48
Met Phe Phe Phe Met Ser Arg Ser Leu Leu Thr Asn Glu Thr Ser Glu
1 5 10 15

ttg gat tta ctg gat caa cgt cct ttc gac cag acc gat ttt gat att 96
Leu Asp Leu Leu Asp Gln Arg Pro Phe Asp Gln Thr Asp Phe Asp Ile
20 25 30

ctg aaa tcc tac gaa gcg gtg gtg gac ggg tta gcg atg ctt att ggc 144
Leu Lys Ser Tyr Glu Ala Val Val Asp Gly Leu Ala Met Leu Ile Gly
35 40 45

tcc cac tgt gaa atc gtt ttg cac tct ttg cag gat cta aaa tgt tca 192
Ser His Cys Glu Ile Val Leu His Ser Leu Gln Asp Leu Lys Cys Ser
50 55 60

gcc att cgc att gct aac ggt gaa cat aca ggc cgg aag att ggt tcg 240
Ala Ile Arg Ile Ala Asn Gly Glu His Thr Gly Arg Lys Ile Gly Ser
65 70 75 80

cca att act gac ctg gcg cta cgt atg ctg cac gat atg acg gga gcg 288
Pro Ile Thr Asp Leu Ala Leu Arg Met Leu His Asp Met Thr Gly Ala
85 90 95

gat agc agc gtt tct aaa tgc tac ttt act cgc gcc aaa agc ggc gta 336
Asp Ser Ser Val Ser Lys Cys Tyr Phe Thr Arg Ala Lys Ser Gly Val
100 105 110

tta atg aag tcc ctg act atc gcg att cgt aac cgc gaa cag cgt gta 384
Leu Met Lys Ser Leu Thr Ile Ala Ile Arg Asn Arg Glu Gln Arg Val
115 120 125

att ggt ctg ctg tgc atc aat atg aat ctt gat gtt ccc ttc tcg cag 432
Ile Gly Leu Leu Cys Ile Asn Met Asn Leu Asp Val Pro Phe Ser Gln
130 135 140

att atg agc acc ttt gtg ccg cca gaa acc ccg gat gtc ggt tca agc 480
Ile Met Ser Thr Phe Val Pro Pro Glu Thr Pro Asp Val Gly Ser Ser
145 150 155 160

gtc aac ttt gcc tct tct gtt gaa gat ctg gtt acc caa acg ctg gag 528
Val Asn Phe Ala Ser Ser Val Glu Asp Leu Val Thr Gln Thr Leu Glu
165 170 175

ttc acc atc gaa gaa gtg aat gcc gat cgc aat gtt tct aat aac gcc 576
Phe Thr Ile Glu Glu Val Asn Ala Asp Arg Asn Val Ser Asn Asn Ala

```
            180                 185                 190

aaa aat cgt cag atc gtg ctg aat ctc tac gag aaa ggg atc ttc gat    624
Lys Asn Arg Gln Ile Val Leu Asn Leu Tyr Glu Lys Gly Ile Phe Asp
        195                 200                 205

att aaa gat gcg atc aac cag gtt gct gac cgc ctg aac atc tcc aaa    672
Ile Lys Asp Ala Ile Asn Gln Val Ala Asp Arg Leu Asn Ile Ser Lys
    210                 215                 220

cac act gtc tat ctc tac atc cgc cag ttc aag agc ggt gat ttc cag    720
His Thr Val Tyr Leu Tyr Ile Arg Gln Phe Lys Ser Gly Asp Phe Gln
225                 230                 235                 240

ggg caa gat aag taa                                                735
Gly Gln Asp Lys  *
```

<210> 123
<211> 255
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(255)

<400> 123
```
atg acc gat aaa atc cgt act ctg caa ggt cgc gtt gtt agc gac aaa    48
Met Thr Asp Lys Ile Arg Thr Leu Gln Gly Arg Val Val Ser Asp Lys
 1               5                  10                  15

atg gag aaa tcc att gtt gtt gct atc gaa cgt ttt gtg aaa cac ccg    96
Met Glu Lys Ser Ile Val Val Ala Ile Glu Arg Phe Val Lys His Pro
            20                  25                  30

atc tac ggt aaa ttc atc aag cgt acg acc aaa ctg cac gta cat gac    144
Ile Tyr Gly Lys Phe Ile Lys Arg Thr Thr Lys Leu His Val His Asp
        35                  40                  45

gag aac aac gaa tgc ggt atc ggt gac gtg gtt gaa atc cgc gaa tgc    192
Glu Asn Asn Glu Cys Gly Ile Gly Asp Val Val Glu Ile Arg Glu Cys
    50                  55                  60

cgt ccg ctg tcc aag act aaa tcc tgg acg ctg gtt cgc gtt gta gag    240
Arg Pro Leu Ser Lys Thr Lys Ser Trp Thr Leu Val Arg Val Val Glu
 65                 70                  75                  80

aaa gcg gtt ctg taa                                                255
Lys Ala Val Leu  *
```

<210> 124
<211> 192
<212> DNA
<213> Escherichia coli

<220>
<221> CDS

<222> (1)...(192)

<400> 124
atg aaa gca aaa gag ctg cgt gag aag agc gtt gaa gag ctg aac acc 48
Met Lys Ala Lys Glu Leu Arg Glu Lys Ser Val Glu Glu Leu Asn Thr
1 5 10 15

gag ctg ctg aac ctg ctg cgt gag cag ttc aac ctg cgt atg cag gct 96
Glu Leu Leu Asn Leu Leu Arg Glu Gln Phe Asn Leu Arg Met Gln Ala
20 25 30

gca agt ggc cag ctg caa cag tct cac ctg ttg aag caa gtg cgt cgc 144
Ala Ser Gly Gln Leu Gln Gln Ser His Leu Leu Lys Gln Val Arg Arg
35 40 45

gat gtc gca cgc gtt aag act tta ctg aac gag aag gcg ggt gcg taa 192
Asp Val Ala Arg Val Lys Thr Leu Leu Asn Glu Lys Ala Gly Ala *
50 55 60

<210> 125
<211> 411
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(411)

<400> 125
atg tta caa cca aag cgt aca aaa ttc cgt aaa atg cac aaa ggc cgt 48
Met Leu Gln Pro Lys Arg Thr Lys Phe Arg Lys Met His Lys Gly Arg
1 5 10 15

aac cgc ggt ctg gcg cag ggt acg gat gtt agc ttc ggc agc ttc ggt 96
Asn Arg Gly Leu Ala Gln Gly Thr Asp Val Ser Phe Gly Ser Phe Gly
20 25 30

ctg aaa gct gtt ggc cgt ggt cgt ctg act gcc cgt cag atc gaa gca 144
Leu Lys Ala Val Gly Arg Gly Arg Leu Thr Ala Arg Gln Ile Glu Ala
35 40 45

gca cgt cgt gct atg acc cgt gca gtt aag cgt caa ggt aag atc tgg 192
Ala Arg Arg Ala Met Thr Arg Ala Val Lys Arg Gln Gly Lys Ile Trp
50 55 60

atc cgt gtg ttc ccg gac aaa ccg atc act gaa aag ccg ctg gca gtg 240
Ile Arg Val Phe Pro Asp Lys Pro Ile Thr Glu Lys Pro Leu Ala Val
65 70 75 80

cgt atg ggt aaa ggt aaa ggt aac gtg gag tat tgg gtt gcc ttg att 288
Arg Met Gly Lys Gly Lys Gly Asn Val Glu Tyr Trp Val Ala Leu Ile
85 90 95

cag ccg ggt aaa gtc ctg tat gaa atg gac ggt gtt ccg gaa gag ctg 336
Gln Pro Gly Lys Val Leu Tyr Glu Met Asp Gly Val Pro Glu Glu Leu
100 105 110

gcc cgt gaa gca ttc aag ctg gca gca gcg aaa ctg ccg att aaa acc 384
Ala Arg Glu Ala Phe Lys Leu Ala Ala Ala Lys Leu Pro Ile Lys Thr

115 120 125

acc ttt gta act aag acg gtg atg taa 411
Thr Phe Val Thr Lys Thr Val Met *
130 135

<210> 126
<211> 702
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(702)

<400> 126
atg ggt cag aaa gta cat cct aat ggt att cgc ctg ggt att gta aaa 48
Met Gly Gln Lys Val His Pro Asn Gly Ile Arg Leu Gly Ile Val Lys
1 5 10 15

cca tgg aac tct acc tgg ttt gcg aac acc aaa gaa ttc gct gac aac 96
Pro Trp Asn Ser Thr Trp Phe Ala Asn Thr Lys Glu Phe Ala Asp Asn
20 25 30

ctg gac agc gat ttt aaa gta cgt cag tac ctg act aag gaa ctg gct 144
Leu Asp Ser Asp Phe Lys Val Arg Gln Tyr Leu Thr Lys Glu Leu Ala
35 40 45

aaa gcg tcc gta tct cgt atc gtt atc gag cgt ccg gct aag agc atc 192
Lys Ala Ser Val Ser Arg Ile Val Ile Glu Arg Pro Ala Lys Ser Ile
50 55 60

cgt gta acc att cac act gct cgc ccg ggt atc gtt atc ggt aaa aaa 240
Arg Val Thr Ile His Thr Ala Arg Pro Gly Ile Val Ile Gly Lys Lys
65 70 75 80

ggt gaa gac gta gaa aaa ctg cgt aag gtc gta gcg gac atc gct ggc 288
Gly Glu Asp Val Glu Lys Leu Arg Lys Val Val Ala Asp Ile Ala Gly
85 90 95

gtt cct gca cag atc aac atc gcc gaa gtt cgt aag cct gaa ctg gac 336
Val Pro Ala Gln Ile Asn Ile Ala Glu Val Arg Lys Pro Glu Leu Asp
100 105 110

gca aaa ctg gtt gct gac agc atc act tct cag ctg gaa cgt cgc gtt 384
Ala Lys Leu Val Ala Asp Ser Ile Thr Ser Gln Leu Glu Arg Arg Val
115 120 125

atg ttc cgt cgt gct atg aag cgt gct gta cag aac gca atg cgt ctg 432
Met Phe Arg Arg Ala Met Lys Arg Ala Val Gln Asn Ala Met Arg Leu
130 135 140

ggc gct aaa ggt att aaa gtt gaa gtt agc ggc cgt ctg ggc ggc gcg 480
Gly Ala Lys Gly Ile Lys Val Glu Val Ser Gly Arg Leu Gly Gly Ala
145 150 155 160

gaa atc gca cgt acc gaa tgg tac cgc gaa ggt cgc gta ccg ctg cac 528
Glu Ile Ala Arg Thr Glu Trp Tyr Arg Glu Gly Arg Val Pro Leu His
165 170 175

```
act ctg cgt gct gac atc gac tac aac acc tct gaa gcg cac acc act      576
Thr Leu Arg Ala Asp Ile Asp Tyr Asn Thr Ser Glu Ala His Thr Thr
            180                 185                 190

tac ggt gta atc ggc gtt aaa gtg tgg atc ttc aaa ggc gag atc ctg      624
Tyr Gly Val Ile Gly Val Lys Val Trp Ile Phe Lys Gly Glu Ile Leu
        195                 200                 205

ggt ggt atg gct gct gtt gaa caa ccg gaa aaa ccg gct gct cag cct      672
Gly Gly Met Ala Ala Val Glu Gln Pro Glu Lys Pro Ala Ala Gln Pro
    210                 215                 220

aaa aag cag cag cgt aaa ggc cgt aaa taa                              702
Lys Lys Gln Gln Arg Lys Gly Arg Lys  *
225                 230


<210> 127
<211> 333
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(333)

<400> 127
atg gaa act atc gct aaa cat cgc cat gct cgt tct tct gct cag aag      48
Met Glu Thr Ile Ala Lys His Arg His Ala Arg Ser Ser Ala Gln Lys
 1               5                  10                  15

gtt cgc ctt gtt gct gac ctg att cgc ggt aag aaa gtg tcg cag gct      96
Val Arg Leu Val Ala Asp Leu Ile Arg Gly Lys Lys Val Ser Gln Ala
            20                  25                  30

ctg gat att ttg acc tac acc aac aag aaa gcg gct gta ctg gtc aag      144
Leu Asp Ile Leu Thr Tyr Thr Asn Lys Lys Ala Ala Val Leu Val Lys
        35                  40                  45

aaa gtt ctg gaa tct gcc att gct aac gct gaa cac aac gat ggc gct      192
Lys Val Leu Glu Ser Ala Ile Ala Asn Ala Glu His Asn Asp Gly Ala
    50                  55                  60

gac att gac gat ctg aaa gtt acg aaa att ttc gta gac gaa ggc ccg      240
Asp Ile Asp Asp Leu Lys Val Thr Lys Ile Phe Val Asp Glu Gly Pro
65                  70                  75                  80

agc atg aag cgc att atg ccg cgt gca aaa ggt cgt gca gat cgc atc      288
Ser Met Lys Arg Ile Met Pro Arg Ala Lys Gly Arg Ala Asp Arg Ile
                85                  90                  95

ctg aag cgc acc agc cac atc act gtg gtt gtg tcc gat cgc tga          333
Leu Lys Arg Thr Ser His Ile Thr Val Val Val Ser Asp Arg  *
            100                 105                 110


<210> 128
<211> 279
<212> DNA
<213> Escherichia coli
```

<220>
<221> CDS
<222> (1)...(279)

<400> 128
atg cca cgt tct ctc aag aaa ggt cct ttt att gac ctg cac ttg ctg 48
Met Pro Arg Ser Leu Lys Lys Gly Pro Phe Ile Asp Leu His Leu Leu
1 5 10 15

aag aag gta gag aaa gcg gtg gaa agc gga gac aag aag ccc ctg cgc 96
Lys Lys Val Glu Lys Ala Val Glu Ser Gly Asp Lys Lys Pro Leu Arg
20 25 30

act tgg tcc cgt cgt tca acg atc ttt cct aac atg atc ggt ttg acc 144
Thr Trp Ser Arg Arg Ser Thr Ile Phe Pro Asn Met Ile Gly Leu Thr
35 40 45

atc gct gtc cat aat ggt cgt cag cac gtt ccg gta ttt gta acc gac 192
Ile Ala Val His Asn Gly Arg Gln His Val Pro Val Phe Val Thr Asp
50 55 60

gaa atg gtt ggt cac aaa ctg ggt gaa ttc gca ccg act cgt act tat 240
Glu Met Val Gly His Lys Leu Gly Glu Phe Ala Pro Thr Arg Thr Tyr
65 70 75 80

cgc ggc cac gct gct gat aaa aaa gcg aag aag aaa taa 279
Arg Gly His Ala Ala Asp Lys Lys Ala Lys Lys Lys *
85 90

<210> 129
<211> 822
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(822)

<400> 129
atg gca gtt gtt aaa tgt aaa ccg aca tct ccg ggt cgt cgc cac gta 48
Met Ala Val Val Lys Cys Lys Pro Thr Ser Pro Gly Arg Arg His Val
1 5 10 15

gtt aaa gtg gtt aac cct gag ctg cac aag ggc aaa cct ttt gct ccg 96
Val Lys Val Val Asn Pro Glu Leu His Lys Gly Lys Pro Phe Ala Pro
20 25 30

ttg ctg gaa aaa aac agc aaa tcc ggt ggt cgt aac aac aat ggc cgt 144
Leu Leu Glu Lys Asn Ser Lys Ser Gly Gly Arg Asn Asn Asn Gly Arg
35 40 45

atc acc act cgt cat atc ggt ggt ggc cac aag cag gct tac cgt att 192
Ile Thr Thr Arg His Ile Gly Gly Gly His Lys Gln Ala Tyr Arg Ile
50 55 60

gtt gac ttc aaa cgc aac aaa gac ggt atc ccg gca gtt gtt gaa cgt 240
Val Asp Phe Lys Arg Asn Lys Asp Gly Ile Pro Ala Val Val Glu Arg
65 70 75 80

ctt gag tac gat ccg aac cgt tcc gcg aac atc gcg ctg gtt ctg tac 288
Leu Glu Tyr Asp Pro Asn Arg Ser Ala Asn Ile Ala Leu Val Leu Tyr
85 90 95

aaa gac ggt gaa cgc cgt tac atc ctg gcc cct aaa ggc ctg aaa gct 336
Lys Asp Gly Glu Arg Arg Tyr Ile Leu Ala Pro Lys Gly Leu Lys Ala
100 105 110

ggc gac cag att cag tct ggc gtt gat gct gca atc aaa cca ggt aac 384
Gly Asp Gln Ile Gln Ser Gly Val Asp Ala Ala Ile Lys Pro Gly Asn
115 120 125

acc ctg ccg atg cgc aac atc ccg gtt ggt tct act gtt cat aac gta 432
Thr Leu Pro Met Arg Asn Ile Pro Val Gly Ser Thr Val His Asn Val
130 135 140

gaa atg aaa cca ggt aaa ggc ggt cag ctg gca cgt tcc gct ggt act 480
Glu Met Lys Pro Gly Lys Gly Gly Gln Leu Ala Arg Ser Ala Gly Thr
145 150 155 160

tac gtt cag atc gtt gct cgt gat ggt gct tat gtc acc ctg cgt ctg 528
Tyr Val Gln Ile Val Ala Arg Asp Gly Ala Tyr Val Thr Leu Arg Leu
165 170 175

cgt tct ggt gaa atg cgt aaa gta gaa gca gac tgc cgt gca act ctg 576
Arg Ser Gly Glu Met Arg Lys Val Glu Ala Asp Cys Arg Ala Thr Leu
180 185 190

ggc gaa gtt ggc aat gct gag cat atg ctg cgc gtt ctg ggt aaa gca 624
Gly Glu Val Gly Asn Ala Glu His Met Leu Arg Val Leu Gly Lys Ala
195 200 205

ggt gct gca cgc tgg cgt ggt gtt cgt ccg acc gtt cgc ggt acc gcg 672
Gly Ala Ala Arg Trp Arg Gly Val Arg Pro Thr Val Arg Gly Thr Ala
210 215 220

atg aac ccg gta gac cac cca cat ggt ggt ggt gaa ggt cgt aac ttt 720
Met Asn Pro Val Asp His Pro His Gly Gly Gly Glu Gly Arg Asn Phe
225 230 235 240

ggt aag cac ccg gta act ccg tgg ggc gtt cag acc aaa ggt aag aag 768
Gly Lys His Pro Val Thr Pro Trp Gly Val Gln Thr Lys Gly Lys Lys
245 250 255

acc cgc agc aac aag cgt act gat aaa ttc atc gta cgt cgc cgt agc 816
Thr Arg Ser Asn Lys Arg Thr Asp Lys Phe Ile Val Arg Arg Arg Ser
260 265 270

aaa taa 822
Lys *

<210> 130
<211> 303
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(303)

<400> 130
atg att cgt gaa gaa cgt ctg ctg aag gtg ctg cgt gca ccg cac gtt 48
Met Ile Arg Glu Glu Arg Leu Leu Lys Val Leu Arg Ala Pro His Val
1 5 10 15

tct gaa aaa gcg tct act gcg atg gaa aaa tcc aac acc atc gta ctc 96
Ser Glu Lys Ala Ser Thr Ala Met Glu Lys Ser Asn Thr Ile Val Leu
20 25 30

aaa gtt gct aaa gac gcg acc aaa gca gaa atc aaa gct gct gtg cag 144
Lys Val Ala Lys Asp Ala Thr Lys Ala Glu Ile Lys Ala Ala Val Gln
35 40 45

aaa ctg ttt gaa gtc gaa gtc gaa gtc gtt aac acc ctg gta gtt aaa 192
Lys Leu Phe Glu Val Glu Val Glu Val Val Asn Thr Leu Val Val Lys
50 55 60

ggg aaa gtt aaa cgt cac gga cag cgt atc ggt cgt cgt agc gac tgg 240
Gly Lys Val Lys Arg His Gly Gln Arg Ile Gly Arg Arg Ser Asp Trp
65 70 75 80

aaa aaa gct tac gtc acc ctg aaa gaa ggc cag aat ctg gac ttc gtt 288
Lys Lys Ala Tyr Val Thr Leu Lys Glu Gly Gln Asn Leu Asp Phe Val
85 90 95

ggc ggc gct gag taa 303
Gly Gly Ala Glu *
100

<210> 131
<211> 606
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(606)

<400> 131
atg gaa tta gta ttg aaa gac gcg cag agc gcg ctg act gtt tcc gaa 48
Met Glu Leu Val Leu Lys Asp Ala Gln Ser Ala Leu Thr Val Ser Glu
1 5 10 15

act acc ttc ggt cgt gat ttc aac gaa gcg ctg gtt cac cag gtt gtt 96
Thr Thr Phe Gly Arg Asp Phe Asn Glu Ala Leu Val His Gln Val Val
20 25 30

gtt gct tat gca gct ggt gct cgt cag ggt act cgt gct cag aag act 144
Val Ala Tyr Ala Ala Gly Ala Arg Gln Gly Thr Arg Ala Gln Lys Thr
35 40 45

cgt gct gaa gta act ggt tcc ggt aaa aaa ccg tgg cgc cag aaa ggc 192
Arg Ala Glu Val Thr Gly Ser Gly Lys Lys Pro Trp Arg Gln Lys Gly
50 55 60

acc ggc cgt gcg cgt tct ggt tct atc aag agc ccg atc tgg cgt tct 240
Thr Gly Arg Ala Arg Ser Gly Ser Ile Lys Ser Pro Ile Trp Arg Ser
65 70 75 80

ggt ggc gtg acc ttt gct gct cgt ccg cag gac cac agt caa aaa gtt 288
Gly Gly Val Thr Phe Ala Ala Arg Pro Gln Asp His Ser Gln Lys Val
85 90 95

aac aag aag atg tac cgc ggc gcg ctg aaa agc atc ctg tcc gaa ctg 336
Asn Lys Lys Met Tyr Arg Gly Ala Leu Lys Ser Ile Leu Ser Glu Leu
100 105 110

gta cgt cag gat cgt ctg atc gtt gtc gag aag ttc tct gta gaa gcg 384
Val Arg Gln Asp Arg Leu Ile Val Val Glu Lys Phe Ser Val Glu Ala
115 120 125

ccg aaa act aag ctg ctg gca cag aaa ctg aaa gac atg gct ctg gaa 432
Pro Lys Thr Lys Leu Leu Ala Gln Lys Leu Lys Asp Met Ala Leu Glu
130 135 140

gat gtg ctg atc atc acc ggt gag ctg gac gaa aac ctg ttc ctg gct 480
Asp Val Leu Ile Ile Thr Gly Glu Leu Asp Glu Asn Leu Phe Leu Ala
145 150 155 160

gcg cgc aac ctg cac aag gtt gac gta cgc gat gca act ggt atc gac 528
Ala Arg Asn Leu His Lys Val Asp Val Arg Asp Ala Thr Gly Ile Asp
165 170 175

ccg gtt agc ctg atc gcc ttc gac aaa gtc gta atg act gct gat gct 576
Pro Val Ser Leu Ile Ala Phe Asp Lys Val Val Met Thr Ala Asp Ala
180 185 190

gtt aag caa gtt gag gag atg ctg gca tga 606
Val Lys Gln Val Glu Glu Met Leu Ala *
195 200

<210> 132
<211> 630
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(630)

<400> 132
atg att ggt tta gtc ggt aaa aaa gtg ggt atg acc cgt atc ttc aca 48
Met Ile Gly Leu Val Gly Lys Lys Val Gly Met Thr Arg Ile Phe Thr
1 5 10 15

gaa gac ggc gtt tct atc cca gta acc gta atc gaa gtt gaa gca aac 96
Glu Asp Gly Val Ser Ile Pro Val Thr Val Ile Glu Val Glu Ala Asn
20 25 30

cgc gtt act cag gtt aaa gac ctg gct aac gat ggc tac cgt gct att 144
Arg Val Thr Gln Val Lys Asp Leu Ala Asn Asp Gly Tyr Arg Ala Ile
35 40 45

cag gtg acc acc ggt gct aaa aaa gct aac cgt gtg acc aag cct gaa 192
Gln Val Thr Thr Gly Ala Lys Lys Ala Asn Arg Val Thr Lys Pro Glu
50 55 60

gct ggc cac ttc gct aaa gct ggc gta gaa gct ggc cgt ggt ctg tgg 240
Ala Gly His Phe Ala Lys Ala Gly Val Glu Ala Gly Arg Gly Leu Trp

65 70 75 80

gaa ttc cgc ctg gct gaa ggc gaa gag ttc act gta ggt cag agc att 288
Glu Phe Arg Leu Ala Glu Gly Glu Glu Phe Thr Val Gly Gln Ser Ile
85 90 95

agc gtt gaa ctg ttt gct gac gtt aaa aaa gtt gac gta act ggc acc 336
Ser Val Glu Leu Phe Ala Asp Val Lys Lys Val Asp Val Thr Gly Thr
100 105 110

tct aaa ggt aaa ggt ttc gca ggt acc gtt aag cgc tgg aac ttc cgt 384
Ser Lys Gly Lys Gly Phe Ala Gly Thr Val Lys Arg Trp Asn Phe Arg
115 120 125

acc cag gac gct act cac ggt aac tcc ttg tct cac cgc gtt ccg ggt 432
Thr Gln Asp Ala Thr His Gly Asn Ser Leu Ser His Arg Val Pro Gly
130 135 140

tct atc ggt cag aac cag act ccg ggc aaa gtg ttc aaa ggc aag aaa 480
Ser Ile Gly Gln Asn Gln Thr Pro Gly Lys Val Phe Lys Gly Lys Lys
145 150 155 160

atg gca ggt cag atg ggt aac gaa cgt gta acc gtt cag agc ctt gac 528
Met Ala Gly Gln Met Gly Asn Glu Arg Val Thr Val Gln Ser Leu Asp
165 170 175

gta gta cgc gtt gac gct gag cgc aac ctg ctg ctg gtt aaa ggt gct 576
Val Val Arg Val Asp Ala Glu Arg Asn Leu Leu Leu Val Lys Gly Ala
180 185 190

gtc ccg ggt gca acc ggt agc gac ctg atc gtt aaa cca gct gtg aag 624
Val Pro Gly Ala Thr Gly Ser Asp Leu Ile Val Lys Pro Ala Val Lys
195 200 205

gcg taa 630
Ala *

<210> 133
<211> 312
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(312)

<400> 133
atg cag aac caa aga atc cgt atc cgc ctg aaa gcg ttt gat cat cgt 48
Met Gln Asn Gln Arg Ile Arg Ile Arg Leu Lys Ala Phe Asp His Arg
1 5 10 15

ctg atc gat caa gca acc gcg gaa atc gtc gag act gcc aag cgc act 96
Leu Ile Asp Gln Ala Thr Ala Glu Ile Val Glu Thr Ala Lys Arg Thr
20 25 30

ggt gcg cag gtc cgt ggt ccg atc ccg ctg ccg aca cgc aaa gag cgc 144
Gly Ala Gln Val Arg Gly Pro Ile Pro Leu Pro Thr Arg Lys Glu Arg
35 40 45

```
ttc act gtt ctg atc tcc ccg cac gtc aac aaa gac gcg cgc gat cag      192
Phe Thr Val Leu Ile Ser Pro His Val Asn Lys Asp Ala Arg Asp Gln
     50                  55                  60

tac gaa atc cgt act cac ttg cgt ctg gtt gac atc gtt gag cca acc      240
Tyr Glu Ile Arg Thr His Leu Arg Leu Val Asp Ile Val Glu Pro Thr
 65                  70                  75                  80

gag aaa acc gtt gat gct ctg atg cgt ctg gat ctg gct gcc ggt gta      288
Glu Lys Thr Val Asp Ala Leu Met Arg Leu Asp Leu Ala Ala Gly Val
                 85                  90                  95

gac gtg cag atc agc ctg ggt taa                                      312
Asp Val Gln Ile Ser Leu Gly  *
            100
```

<210> 134
<211> 315
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(315)

<400> 134

```
atg atc cgc aaa gcc ttt gtc atg cag gta aac ccc gac gcc cac gaa      48
Met Ile Arg Lys Ala Phe Val Met Gln Val Asn Pro Asp Ala His Glu
 1               5                  10                  15

gag tat cag cgt cgg cat aat ccc atc tgg cca gaa ctg gaa gca gtg      96
Glu Tyr Gln Arg Arg His Asn Pro Ile Trp Pro Glu Leu Glu Ala Val
            20                  25                  30

ctg aaa tct cac ggt gcg cat aac tac gcc atc tat ctc gac aaa gcg      144
Leu Lys Ser His Gly Ala His Asn Tyr Ala Ile Tyr Leu Asp Lys Ala
        35                  40                  45

cgt aat ctg ctg ttt gcc atg gta gag att gaa tct gaa gaa cgc tgg      192
Arg Asn Leu Leu Phe Ala Met Val Glu Ile Glu Ser Glu Glu Arg Trp
     50                  55                  60

aat gcg gtt gcc agc act gat gtt tgc caa cgt tgg tgg aaa tat atg      240
Asn Ala Val Ala Ser Thr Asp Val Cys Gln Arg Trp Trp Lys Tyr Met
 65                  70                  75                  80

acc gat gtt atg ccc gct aac ccg gat aac agc ccg gtg agt agc gag      288
Thr Asp Val Met Pro Ala Asn Pro Asp Asn Ser Pro Val Ser Ser Glu
                 85                  90                  95

ctg caa gaa gtg ttt tac ctg ccg taa                                  315
Leu Gln Glu Val Phe Tyr Leu Pro  *
            100
```

<210> 135
<211> 864
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(864)

<400> 135
atg atc cgc agt atg acc gcc tac gcc cgg cgt gaa atc aag ggt gaa 48
Met Ile Arg Ser Met Thr Ala Tyr Ala Arg Arg Glu Ile Lys Gly Glu
1 5 10 15

tgg ggg agc gca acc tgg gaa atg cgc tcg gta aac cag cgt tat ctg 96
Trp Gly Ser Ala Thr Trp Glu Met Arg Ser Val Asn Gln Arg Tyr Leu
20 25 30

gaa act tac ttt cgt ctg ccg gag cag ttc cgt agc ctt gaa cct gtc 144
Glu Thr Tyr Phe Arg Leu Pro Glu Gln Phe Arg Ser Leu Glu Pro Val
35 40 45

gtt cgc gag cgt att cgt tct cgc ctg acg cgc ggt aaa gtg gaa tgt 192
Val Arg Glu Arg Ile Arg Ser Arg Leu Thr Arg Gly Lys Val Glu Cys
50 55 60

acc ctg cgc tat gag cca gat gtt agc gcg caa ggt gag ctg atc ctc 240
Thr Leu Arg Tyr Glu Pro Asp Val Ser Ala Gln Gly Glu Leu Ile Leu
65 70 75 80

aac gaa aaa ctg gct aaa cag ctg gta act gcc gcg aac tgg gta aaa 288
Asn Glu Lys Leu Ala Lys Gln Leu Val Thr Ala Ala Asn Trp Val Lys
85 90 95

atg cag agt gac gaa ggg gaa atc aac ccg gtt gat att cta cgc tgg 336
Met Gln Ser Asp Glu Gly Glu Ile Asn Pro Val Asp Ile Leu Arg Trp
100 105 110

ccg ggc gtg atg gca gcc cag gag cag gat ctt gac gcc att gcc gct 384
Pro Gly Val Met Ala Ala Gln Glu Gln Asp Leu Asp Ala Ile Ala Ala
115 120 125

gaa att ctc gcg gcg ctg gat ggt acg ctg gac gac ttt att gtc gcg 432
Glu Ile Leu Ala Ala Leu Asp Gly Thr Leu Asp Asp Phe Ile Val Ala
130 135 140

cgc gaa acc gaa ggt cag gca ctg aaa gca ttg atc gag cag cgt ctg 480
Arg Glu Thr Glu Gly Gln Ala Leu Lys Ala Leu Ile Glu Gln Arg Leu
145 150 155 160

gaa ggc gtc acc gcc gaa gtg gtc aaa gtc cgc tcc cat atg ccg gaa 528
Glu Gly Val Thr Ala Glu Val Val Lys Val Arg Ser His Met Pro Glu
165 170 175

atc ctg caa tgg cag cgt gag cgt ctg gtc gcg aag ctg gaa gat gct 576
Ile Leu Gln Trp Gln Arg Glu Arg Leu Val Ala Lys Leu Glu Asp Ala
180 185 190

cag gtg caa ctg gaa aac aac cgt ctg gag cag gaa ctg gtt ctg ctg 624
Gln Val Gln Leu Glu Asn Asn Arg Leu Glu Gln Glu Leu Val Leu Leu
195 200 205

gca caa cga att gac gtt gcc gaa gaa ctg gat cgc ctc gaa gcg cat 672
Ala Gln Arg Ile Asp Val Ala Glu Glu Leu Asp Arg Leu Glu Ala His
210 215 220

gtc aaa gag acc tac aac att ctg aag aaa aaa gaa gcg gtt ggt cgt 720
Val Lys Glu Thr Tyr Asn Ile Leu Lys Lys Lys Glu Ala Val Gly Arg
225 230 235 240

cgt ctg gat ttt atg atg cag gag ttc aac cgc gag tcg aac act ctt 768
Arg Leu Asp Phe Met Met Gln Glu Phe Asn Arg Glu Ser Asn Thr Leu
245 250 255

gcg tcg aag tct atc aat gcc gaa gtg aca aac tcc gcc atc gag ctg 816
Ala Ser Lys Ser Ile Asn Ala Glu Val Thr Asn Ser Ala Ile Glu Leu
260 265 270

aaa gtg ttg att gag cag atg cgc gag cag att cag aac atc gaa taa 864
Lys Val Leu Ile Glu Gln Met Arg Glu Gln Ile Gln Asn Ile Glu *
275 280 285

<210> 136
<211> 492
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(492)

<400> 136
atg agc aca gaa aca att gaa ata ttc aat aat agt gat gaa tgg gca 48
Met Ser Thr Glu Thr Ile Glu Ile Phe Asn Asn Ser Asp Glu Trp Ala
1 5 10 15

aat caa cta aaa cac gca tta tcg aaa gga gaa aat ctg gca tta cta 96
Asn Gln Leu Lys His Ala Leu Ser Lys Gly Glu Asn Leu Ala Leu Leu
20 25 30

cat ggt tta act cct gat atc ctt gat aga ata tat gca tat gca ttc 144
His Gly Leu Thr Pro Asp Ile Leu Asp Arg Ile Tyr Ala Tyr Ala Phe
35 40 45

gac tac cat gaa aag ggt aat ata acg gac gca gaa att tat tat aaa 192
Asp Tyr His Glu Lys Gly Asn Ile Thr Asp Ala Glu Ile Tyr Tyr Lys
50 55 60

ttc ttg tgc att tat gcg ttc gaa aat cat gag tat cta aaa gat ttt 240
Phe Leu Cys Ile Tyr Ala Phe Glu Asn His Glu Tyr Leu Lys Asp Phe
65 70 75 80

gca tca gta tgt cag ccc aaa aag aaa tat caa caa gca tat gac ctt 288
Ala Ser Val Cys Gln Pro Lys Lys Lys Tyr Gln Gln Ala Tyr Asp Leu
85 90 95

tac aaa cta agt tac aat tac ttc ccg tat gat gac tat tca gtt att 336
Tyr Lys Leu Ser Tyr Asn Tyr Phe Pro Tyr Asp Asp Tyr Ser Val Ile
100 105 110

tat cgt atg ggt caa tgt cag att ggg gct aaa aat atc gat aac gca 384
Tyr Arg Met Gly Gln Cys Gln Ile Gly Ala Lys Asn Ile Asp Asn Ala
115 120 125

atg caa tgt ttc tat cac att att aac aat tgt gag gat gat agt gtt 432

```
Met Gln Cys Phe Tyr His Ile Ile Asn Asn Cys Glu Asp Asp Ser Val
    130                 135                 140

aag agt aaa gcg cag gca tat att gaa ctc tta aac gat aat tca gaa     480
Lys Ser Lys Ala Gln Ala Tyr Ile Glu Leu Leu Asn Asp Asn Ser Glu
145                 150                 155                 160

gat aat ggc taa                                                     492
Asp Asn Gly  *



&lt;210&gt; 137
&lt;211&gt; 1947
&lt;212&gt; DNA
&lt;213&gt; Escherichia coli

&lt;220&gt;
&lt;221&gt; CDS
&lt;222&gt; (1)...(1947)

&lt;400&gt; 137
atg aat att tta ggt ttt ttc cag cga ctc ggt agg gcg tta cag ctc     48
Met Asn Ile Leu Gly Phe Phe Gln Arg Leu Gly Arg Ala Leu Gln Leu
 1               5                  10                  15

cct atc gcg gtg ctg ccg gtg gcg gca ctg ttg ctg cga ttc ggt cag     96
Pro Ile Ala Val Leu Pro Val Ala Ala Leu Leu Leu Arg Phe Gly Gln
            20                  25                  30

cca gat tta ctt aac gtt gcg ttt att gcc cag gcg ggc ggt gcg att     144
Pro Asp Leu Leu Asn Val Ala Phe Ile Ala Gln Ala Gly Gly Ala Ile
        35                  40                  45

ttt gat aac ctc gca tta atc ttc gcc atc ggt gtg gca tcc agc tgg     192
Phe Asp Asn Leu Ala Leu Ile Phe Ala Ile Gly Val Ala Ser Ser Trp
    50                  55                  60

tcg aaa gac agc gct ggt gcg gcg gcg ctg gcg ggt gcg gta ggt tac     240
Ser Lys Asp Ser Ala Gly Ala Ala Ala Leu Ala Gly Ala Val Gly Tyr
 65                 70                  75                  80

ttt gtg tta acc aaa gcg atg gtg acc atc aac cca gaa att aac atg     288
Phe Val Leu Thr Lys Ala Met Val Thr Ile Asn Pro Glu Ile Asn Met
                85                  90                  95

ggt gta ctg gcg ggt atc att acc ggt ctg gtt ggt ggc gca gcc tat     336
Gly Val Leu Ala Gly Ile Ile Thr Gly Leu Val Gly Gly Ala Ala Tyr
            100                 105                 110

aac cgt tgg tcc gat att aaa ctg ccg gac ttc ctg agc ttc ttc ggc     384
Asn Arg Trp Ser Asp Ile Lys Leu Pro Asp Phe Leu Ser Phe Phe Gly
        115                 120                 125

ggc aaa cgc ttt gtg ccg att gcc acc gga ttc ttc tgc ctg gtg ctg     432
Gly Lys Arg Phe Val Pro Ile Ala Thr Gly Phe Phe Cys Leu Val Leu
    130                 135                 140

gcg gcc att ttt ggt tac gtc tgg ccg ccg gta cag cac gct atc cat     480
Ala Ala Ile Phe Gly Tyr Val Trp Pro Pro Val Gln His Ala Ile His
145                 150                 155                 160
```

```
gca ggc ggc gag tgg atc gtt tct gcg ggc gcg ctg ggt tcc ggt atc    528
Ala Gly Gly Glu Trp Ile Val Ser Ala Gly Ala Leu Gly Ser Gly Ile
                165                 170                 175

ttt ggt ttc atc aac cgt ctg ctg atc cca acc ggt ctg cat cag gta    576
Phe Gly Phe Ile Asn Arg Leu Leu Ile Pro Thr Gly Leu His Gln Val
            180                 185                 190

ctg aac acc atc gcc tgg ttc cag att ggt gaa ttc acc aac gcg gcg    624
Leu Asn Thr Ile Ala Trp Phe Gln Ile Gly Glu Phe Thr Asn Ala Ala
        195                 200                 205

ggt acg gtt ttc cac ggt gac att aac cgc ttc tat gcc ggt gac ggc    672
Gly Thr Val Phe His Gly Asp Ile Asn Arg Phe Tyr Ala Gly Asp Gly
    210                 215                 220

acc gcg ggg atg ttc atg tcc ggc ttc ttc ccg atc atg atg ttc ggt    720
Thr Ala Gly Met Phe Met Ser Gly Phe Phe Pro Ile Met Met Phe Gly
225                 230                 235                 240

ctg ccg ggt gcg gcg ctg gcg atg tac ttc gca gca ccg aaa gag cgt    768
Leu Pro Gly Ala Ala Leu Ala Met Tyr Phe Ala Ala Pro Lys Glu Arg
                245                 250                 255

cgt ccg atg gtt ggc ggt atg ctg ctt tct gtt gct gtt act gcg ttc    816
Arg Pro Met Val Gly Gly Met Leu Leu Ser Val Ala Val Thr Ala Phe
            260                 265                 270

ctg acc ggt gtg act gag ccg ctg gaa ttc ctg ttc atg ttc ctt gct    864
Leu Thr Gly Val Thr Glu Pro Leu Glu Phe Leu Phe Met Phe Leu Ala
        275                 280                 285

ccg ctg ctg tac ctc ctg cac gca ctg ctg acc ggt atc agc ctg ttt    912
Pro Leu Leu Tyr Leu Leu His Ala Leu Leu Thr Gly Ile Ser Leu Phe
    290                 295                 300

gtg gca acg ctg ctg ggt atc cac gcg ggc ttc tct ttc tct gcg ggg    960
Val Ala Thr Leu Leu Gly Ile His Ala Gly Phe Ser Phe Ser Ala Gly
305                 310                 315                 320

gct atc gac tac gcg ttg atg tat aac ctg ccg gcc gcc agc cag aac    1008
Ala Ile Asp Tyr Ala Leu Met Tyr Asn Leu Pro Ala Ala Ser Gln Asn
                325                 330                 335

gtc tgg atg ctg ctg gtg atg ggc gtt atc ttc ttc gct atc tac ttc    1056
Val Trp Met Leu Leu Val Met Gly Val Ile Phe Phe Ala Ile Tyr Phe
            340                 345                 350

gtg gtg ttc agt ttg gtt atc cgc atg ttc aac ctg aaa acg ccg ggt    1104
Val Val Phe Ser Leu Val Ile Arg Met Phe Asn Leu Lys Thr Pro Gly
        355                 360                 365

cgt gaa gat aaa gaa gac gag atc gtt act gaa gaa gcc aac agc aac    1152
Arg Glu Asp Lys Glu Asp Glu Ile Val Thr Glu Glu Ala Asn Ser Asn
    370                 375                 380

act gaa gaa ggt ctg act caa ctg gca acc aac tat att gct gcg gtt    1200
Thr Glu Glu Gly Leu Thr Gln Leu Ala Thr Asn Tyr Ile Ala Ala Val
385                 390                 395                 400
```

ggc ggc act gac aac ctg aaa gcg att gac gcc tgt atc acc cgt ctg 1248
Gly Gly Thr Asp Asn Leu Lys Ala Ile Asp Ala Cys Ile Thr Arg Leu
405 410 415

cgc ctt aca gtg gct gac tct gcc cgc gtt aac gat acg atg tgt aaa 1296
Arg Leu Thr Val Ala Asp Ser Ala Arg Val Asn Asp Thr Met Cys Lys
420 425 430

cgt ctg ggt gct tct ggg gta gtg aaa ctg aac aaa cag act att cag 1344
Arg Leu Gly Ala Ser Gly Val Val Lys Leu Asn Lys Gln Thr Ile Gln
435 440 445

gtg att gtt ggc gcg aaa gca gaa tcc atc ggc gat gcg atg aag aaa 1392
Val Ile Val Gly Ala Lys Ala Glu Ser Ile Gly Asp Ala Met Lys Lys
450 455 460

gtc gtt gcc cgt ggt ccg gta gcc gct gcg tca gct gaa gca act ccg 1440
Val Val Ala Arg Gly Pro Val Ala Ala Ala Ser Ala Glu Ala Thr Pro
465 470 475 480

gca act gcc gcg cct gta gca aaa ccg cag gct gta cca aac gcg gta 1488
Ala Thr Ala Ala Pro Val Ala Lys Pro Gln Ala Val Pro Asn Ala Val
485 490 495

tct atc gcg gag ctg gta tcg ccg att acc ggt gat gtc gtg gca ctg 1536
Ser Ile Ala Glu Leu Val Ser Pro Ile Thr Gly Asp Val Val Ala Leu
500 505 510

gat cag gtt cct gac gaa gca ttc gcc agc aaa gcg gtg ggt gac ggt 1584
Asp Gln Val Pro Asp Glu Ala Phe Ala Ser Lys Ala Val Gly Asp Gly
515 520 525

gtg gcg gtg aaa ccg aca gat aaa atc gtc gta tca cca gcc gca ggg 1632
Val Ala Val Lys Pro Thr Asp Lys Ile Val Val Ser Pro Ala Ala Gly
530 535 540

aca atc gtg aaa atc ttc aac acc aac cac gcg ttc tgc ctg gaa acc 1680
Thr Ile Val Lys Ile Phe Asn Thr Asn His Ala Phe Cys Leu Glu Thr
545 550 555 560

gaa aaa ggc gcg gag atc gtc gtc cat atg ggt atc gac acc gta gcg 1728
Glu Lys Gly Ala Glu Ile Val Val His Met Gly Ile Asp Thr Val Ala
565 570 575

ctg gaa ggt aaa ggc ttt aaa cgt ctg gtg gaa gag ggt gcg cag gta 1776
Leu Glu Gly Lys Gly Phe Lys Arg Leu Val Glu Glu Gly Ala Gln Val
580 585 590

agc gca ggg caa ccg att ctg gaa atg gat ctg gat tac ctg aac gct 1824
Ser Ala Gly Gln Pro Ile Leu Glu Met Asp Leu Asp Tyr Leu Asn Ala
595 600 605

aac gcc cgc tcg atg att agc ccg gtg gtt tgc agc aat atc gac gat 1872
Asn Ala Arg Ser Met Ile Ser Pro Val Val Cys Ser Asn Ile Asp Asp
610 615 620

ttc agt ggc ttg atc att aaa gct cag ggc cat att gtg gcg ggt caa 1920
Phe Ser Gly Leu Ile Ile Lys Ala Gln Gly His Ile Val Ala Gly Gln
625 630 635 640

aca ccg ctg tat gaa atc aaa aag taa 1947

Thr Pro Leu Tyr Glu Ile Lys Lys *
645

<210> 138
<211> 1239
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1239)

<400> 138
atg aag act atc ttc agg tac att ctt ttt tta gca ctg tat tct tgt 48
Met Lys Thr Ile Phe Arg Tyr Ile Leu Phe Leu Ala Leu Tyr Ser Cys
1 5 10 15

tgt aat aca gtc agt gca tat aca agt ttt att gtg gga aat aat gca 96
Cys Asn Thr Val Ser Ala Tyr Thr Ser Phe Ile Val Gly Asn Asn Ala
20 25 30

gga gtt gat aac tat cga ggc ccc tcc act gcc gca cag atg acc ttt 144
Gly Val Asp Asn Tyr Arg Gly Pro Ser Thr Ala Ala Gln Met Thr Phe
35 40 45

aat tac aca tca aca gca agc aac ttg gtt ttt tat aaa ccc acg cag 192
Asn Tyr Thr Ser Thr Ala Ser Asn Leu Val Phe Tyr Lys Pro Thr Gln
50 55 60

ctc ggc ccg act ggg gta aaa atg tac tgg tca tac ctg gat aca ggt 240
Leu Gly Pro Thr Gly Val Lys Met Tyr Trp Ser Tyr Leu Asp Thr Gly
65 70 75 80

acc ggt ggt ggt att ctt tac tgc aat aca tct ggc aga gcg aat cct 288
Thr Gly Gly Gly Ile Leu Tyr Cys Asn Thr Ser Gly Arg Ala Asn Pro
85 90 95

ggt cca ata act att gaa aat gcc atg gtc tat tca ggt aaa gat tat 336
Gly Pro Ile Thr Ile Glu Asn Ala Met Val Tyr Ser Gly Lys Asp Tyr
100 105 110

ggc gga cat aaa cta ttt aat aca tct gtt cct ggt ctg tat tac acc 384
Gly Gly His Lys Leu Phe Asn Thr Ser Val Pro Gly Leu Tyr Tyr Thr
115 120 125

atg tta ata tca agg gtc tgg tct gca tac gat aca ata act gac att 432
Met Leu Ile Ser Arg Val Trp Ser Ala Tyr Asp Thr Ile Thr Asp Ile
130 135 140

caa tcg cca gga atc tat atc gga gat cct tcc aac caa gaa ttt ttc 480
Gln Ser Pro Gly Ile Tyr Ile Gly Asp Pro Ser Asn Gln Glu Phe Phe
145 150 155 160

ttt tcc gtc aca gac agc gat cta caa act aag ggt tgc aac aaa gca 528
Phe Ser Val Thr Asp Ser Asp Leu Gln Thr Lys Gly Cys Asn Lys Ala
165 170 175

gac gac tac gat aag ttt tgg gct att ggt ggt ata gta cac aac ata 576
Asp Asp Tyr Asp Lys Phe Trp Ala Ile Gly Gly Ile Val His Asn Ile
180 185 190

```
act gtt gaa ttt tat aca gat act aat ttc gat cct aca ctt aat cag     624
Thr Val Glu Phe Tyr Thr Asp Thr Asn Phe Asp Pro Thr Leu Asn Gln
        195                 200                 205

caa gtc cag tta tcc agt tca tca aat tac ctt tat tcg ttt aag gcc     672
Gln Val Gln Leu Ser Ser Ser Ser Asn Tyr Leu Tyr Ser Phe Lys Ala
    210                 215                 220

tac agt cct ggt aca aaa gtt gta gat cac agc aac cac att tat gtc     720
Tyr Ser Pro Gly Thr Lys Val Val Asp His Ser Asn His Ile Tyr Val
225                 230                 235                 240

aat ttc aca ctg aat aat gtc aaa tta acc tta cca aca tgt ttt acc     768
Asn Phe Thr Leu Asn Asn Val Lys Leu Thr Leu Pro Thr Cys Phe Thr
                245                 250                 255

tcc ata ctt acc ggg cca tca gtc aat ggt tca acg gtt aga atg gga     816
Ser Ile Leu Thr Gly Pro Ser Val Asn Gly Ser Thr Val Arg Met Gly
            260                 265                 270

gaa tat agc tct ggg aca att aaa aat ggc gct tca cct gtt ccc ttt     864
Glu Tyr Ser Ser Gly Thr Ile Lys Asn Gly Ala Ser Pro Val Pro Phe
        275                 280                 285

gat atc tcg ctt cag aat tgt att cgt gtg cgt aat att gaa aca aaa     912
Asp Ile Ser Leu Gln Asn Cys Ile Arg Val Arg Asn Ile Glu Thr Lys
    290                 295                 300

ctc gtc act gga aaa gta ggt act caa aac aca caa ctg ctt ggt aat     960
Leu Val Thr Gly Lys Val Gly Thr Gln Asn Thr Gln Leu Leu Gly Asn
305                 310                 315                 320

acg ctt act gga agc act gcc gcc aaa ggc gtt ggc gta ctc att gaa     1008
Thr Leu Thr Gly Ser Thr Ala Ala Lys Gly Val Gly Val Leu Ile Glu
                325                 330                 335

ggt tta gca acc agt aaa aat ccg cta atg aca tta aaa cct aat gat     1056
Gly Leu Ala Thr Ser Lys Asn Pro Leu Met Thr Leu Lys Pro Asn Asp
            340                 345                 350

acg aat tct gtt tat ata gat tat gaa acc gaa gat gac acc tcc gat     1104
Thr Asn Ser Val Tyr Ile Asp Tyr Glu Thr Glu Asp Asp Thr Ser Asp
        355                 360                 365

ggg gtt tac cct aat caa ggt aac ggc aca tca cag cct ctt cat ttc     1152
Gly Val Tyr Pro Asn Gln Gly Asn Gly Thr Ser Gln Pro Leu His Phe
    370                 375                 380

cag gca aca tta aag caa gac ggg aat att gct atc gaa ccc ggc gaa     1200
Gln Ala Thr Leu Lys Gln Asp Gly Asn Ile Ala Ile Glu Pro Gly Glu
385                 390                 395                 400

ttt aaa gcc acc agt act ttc cag gta acc tat ccc tga                 1239
Phe Lys Ala Thr Ser Thr Phe Gln Val Thr Tyr Pro  *
                405                 410
```

<210> 139
<211> 597
<212> DNA

<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(597)

<400> 139
```
atg cat ccc act caa cgt aag ctg atg aag aga ata att ctg ttt ctg    48
Met His Pro Thr Gln Arg Lys Leu Met Lys Arg Ile Ile Leu Phe Leu
 1               5                  10                  15

tca tta ctg ttt tgc atc gcc tgt cca gcc att gct gga cag gat att    96
Ser Leu Leu Phe Cys Ile Ala Cys Pro Ala Ile Ala Gly Gln Asp Ile
            20                  25                  30

gac ctt gtt gcc aat gta aaa aac agc acc tgc aaa agc gga atc agt   144
Asp Leu Val Ala Asn Val Lys Asn Ser Thr Cys Lys Ser Gly Ile Ser
        35                  40                  45

aac cag ggt aat att gat ctt ggc gtc gtt ggg gtg gga tat ttt tca   192
Asn Gln Gly Asn Ile Asp Leu Gly Val Val Gly Val Gly Tyr Phe Ser
    50                  55                  60

ggt aat gtt act cct gaa agt tat caa cca ggt gga aaa gag ttc act   240
Gly Asn Val Thr Pro Glu Ser Tyr Gln Pro Gly Gly Lys Glu Phe Thr
65                  70                  75                  80

atc act gta tcc gac tgt gca tta cag gga act ggc gat gtg cta aat   288
Ile Thr Val Ser Asp Cys Ala Leu Gln Gly Thr Gly Asp Val Leu Asn
                85                  90                  95

cag tta cat att gat ttt aga gcc ctt agc ggt gtc atg gct gct ggc   336
Gln Leu His Ile Asp Phe Arg Ala Leu Ser Gly Val Met Ala Ala Gly
            100                 105                 110

tct agg caa ata ttt gct aat gaa att tcg tca gga gca agt aat gta   384
Ser Arg Gln Ile Phe Ala Asn Glu Ile Ser Ser Gly Ala Ser Asn Val
        115                 120                 125

gga gta gtt ata ttt tct act cag gat tcg gcg aat aca ttc aat gtt   432
Gly Val Val Ile Phe Ser Thr Gln Asp Ser Ala Asn Thr Phe Asn Val
    130                 135                 140

ctt aat gct tca ggc gga tct cgt tcc gtt tat cca gta atg tcg gat   480
Leu Asn Ala Ser Gly Gly Ser Arg Ser Val Tyr Pro Val Met Ser Asp
145                 150                 155                 160

gac atg aat ggt tca tcc tgg aaa ttt agc acc cga atg caa aaa atc   528
Asp Met Asn Gly Ser Ser Trp Lys Phe Ser Thr Arg Met Gln Lys Ile
                165                 170                 175

gat cct gca ttg agt gtt aca tct ggt caa ctt atg agc cat gtg tta   576
Asp Pro Ala Leu Ser Val Thr Ser Gly Gln Leu Met Ser His Val Leu
            180                 185                 190

gtg gat att tac tac gaa taa                                       597
Val Asp Ile Tyr Tyr Glu  *
        195
```

<210> 140

<211> 606
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(606)

<400> 140

```
atg atg aca ttt aaa aat tta cgt tat gga tta tcc agc agt gtc gtt      48
Met Met Thr Phe Lys Asn Leu Arg Tyr Gly Leu Ser Ser Ser Val Val
 1               5                  10                  15

ttg gct gcc tca ttg ttc agc gta ctc tct tat gcg gca aca gac agt      96
Leu Ala Ala Ser Leu Phe Ser Val Leu Ser Tyr Ala Ala Thr Asp Ser
            20                  25                  30

att gga ctg acc gtt att act act gta gaa atg ggt act tgt acc gct      144
Ile Gly Leu Thr Val Ile Thr Thr Val Glu Met Gly Thr Cys Thr Ala
        35                  40                  45

aca tta gta aat gac tct gat cag gac att tct gtt gtt gat ttt ggt      192
Thr Leu Val Asn Asp Ser Asp Gln Asp Ile Ser Val Val Asp Phe Gly
    50                  55                  60

gat gta tat att tct gaa atc aat gcc aag acc aaa gta aaa aca ttc      240
Asp Val Tyr Ile Ser Glu Ile Asn Ala Lys Thr Lys Val Lys Thr Phe
 65                  70                  75                  80

aaa ctc aaa ttc aaa gac tgt gcg ggt atc ccc aat aaa aaa gcg caa      288
Lys Leu Lys Phe Lys Asp Cys Ala Gly Ile Pro Asn Lys Lys Ala Gln
                 85                  90                  95

ata aaa tta acc aag cga gcc aca tgc gag gga act gct aat gac ggt      336
Ile Lys Leu Thr Lys Arg Ala Thr Cys Glu Gly Thr Ala Asn Asp Gly
            100                 105                 110

gcg ggg ttt gca aat ggt tcc aca gcc gca gat aaa gca agt gct gtc      384
Ala Gly Phe Ala Asn Gly Ser Thr Ala Ala Asp Lys Ala Ser Ala Val
        115                 120                 125

gcc gtt gaa gtc tgg agc act gta act ccg gca aca ggg agt gca aca      432
Ala Val Glu Val Trp Ser Thr Val Thr Pro Ala Thr Gly Ser Ala Thr
    130                 135                 140

caa ttt agc tgt gta aca cca gca tca caa gag gta aca atc tcc act      480
Gln Phe Ser Cys Val Thr Pro Ala Ser Gln Glu Val Thr Ile Ser Thr
145                 150                 155                 160

gca gcc aat gcg gtc gtt tat tat ccg atg agt gca cgc ctg gtc gtg      528
Ala Ala Asn Ala Val Val Tyr Tyr Pro Met Ser Ala Arg Leu Val Val
                165                 170                 175

gaa aaa aat aaa acc gta aac aat gtc act gcg ggt aag ttt tct gca      576
Glu Lys Asn Lys Thr Val Asn Asn Val Thr Ala Gly Lys Phe Ser Ala
            180                 185                 190

cca gcc aca ttt aca gta acc tat aac taa                              606
Pro Ala Thr Phe Thr Val Thr Tyr Asn  *
        195                 200
```

<210> 141
<211> 612
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(612)

<400> 141
atg gaa ttc ggt gtc aga ttc agt aac tac aag gga cgt caa atg ata 48
Met Glu Phe Gly Val Arg Phe Ser Asn Tyr Lys Gly Arg Gln Met Ile
1 5 10 15

aaa aca acg cca cat aaa ata gtg ata ctg atg gga ata tta tta tca 96
Lys Thr Thr Pro His Lys Ile Val Ile Leu Met Gly Ile Leu Leu Ser
20 25 30

ccc tca gta ttt gca acg gat att aat gta gag ttt aca gcc act gtc 144
Pro Ser Val Phe Ala Thr Asp Ile Asn Val Glu Phe Thr Ala Thr Val
35 40 45

aaa gcg aca acc tgt aac atc aca ctt act ggt aat aac gtc acg aat 192
Lys Ala Thr Thr Cys Asn Ile Thr Leu Thr Gly Asn Asn Val Thr Asn
50 55 60

gat ggc aat aat aac tac aca ttg aga atc cct aag atg ggt ctg gat 240
Asp Gly Asn Asn Asn Tyr Thr Leu Arg Ile Pro Lys Met Gly Leu Asp
65 70 75 80

aag atc gcg aat aaa acg aca gaa tct cag gct gat ttt aaa ctg gtt 288
Lys Ile Ala Asn Lys Thr Thr Glu Ser Gln Ala Asp Phe Lys Leu Val
85 90 95

gcc agt ggg tgc agc agt ggc atc agt tgg att gat acc act ctg acc 336
Ala Ser Gly Cys Ser Ser Gly Ile Ser Trp Ile Asp Thr Thr Leu Thr
100 105 110

gga aat gca tca tca agc tca cct aag ctt att ata ccg cag tct ggt 384
Gly Asn Ala Ser Ser Ser Ser Pro Lys Leu Ile Ile Pro Gln Ser Gly
115 120 125

gat tca tct tcg acg aca agt aat atc ggt atg ggt ttc aaa aaa cgg 432
Asp Ser Ser Ser Thr Thr Ser Asn Ile Gly Met Gly Phe Lys Lys Arg
130 135 140

act act gat gat gcc act ttc ctt aaa cct aac agt gcg gaa aag ata 480
Thr Thr Asp Asp Ala Thr Phe Leu Lys Pro Asn Ser Ala Glu Lys Ile
145 150 155 160

cgc tgg agc aca gac gag atg cag ccc gat aag ggt ctt gaa atg acc 528
Arg Trp Ser Thr Asp Glu Met Gln Pro Asp Lys Gly Leu Glu Met Thr
165 170 175

gtt gcg cta cgt gaa aca gat gca ggg caa ggc gta ccg ggg aat ttt 576
Val Ala Leu Arg Glu Thr Asp Ala Gly Gln Gly Val Pro Gly Asn Phe
180 185 190

cgt gca ctg gcg acg ttt aat ttc atc tat caa taa 612
Arg Ala Leu Ala Thr Phe Asn Phe Ile Tyr Gln *

195 200

<210> 142
<211> 420
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(420)

<400> 142
atg gca atg act tac cac ctg gac gtc gtc agc gca gag caa caa atg 48
Met Ala Met Thr Tyr His Leu Asp Val Val Ser Ala Glu Gln Gln Met
1 5 10 15

ttc tct ggt ctg gtc gag aaa atc cag gta acg ggt agc gaa ggt gaa 96
Phe Ser Gly Leu Val Glu Lys Ile Gln Val Thr Gly Ser Glu Gly Glu
20 25 30

ctg ggg atc tac cct ggc cac gca ccg ctg ctc acc gcc att aag cct 144
Leu Gly Ile Tyr Pro Gly His Ala Pro Leu Leu Thr Ala Ile Lys Pro
35 40 45

ggt atg att cgc atc gtg aaa cag cac ggt cac gaa gag ttt atc tat 192
Gly Met Ile Arg Ile Val Lys Gln His Gly His Glu Glu Phe Ile Tyr
50 55 60

ctg tct ggc ggc att ctt gaa gtg cag cct ggc aac gtg acc gtt ctg 240
Leu Ser Gly Gly Ile Leu Glu Val Gln Pro Gly Asn Val Thr Val Leu
65 70 75 80

gcc gac acc gca att cgc ggc cag gat ctc gac gaa gcg cga gcc atg 288
Ala Asp Thr Ala Ile Arg Gly Gln Asp Leu Asp Glu Ala Arg Ala Met
85 90 95

gaa gcg aaa cgt aag gct gaa gag cac att agc agc tct cac ggc gac 336
Glu Ala Lys Arg Lys Ala Glu Glu His Ile Ser Ser Ser His Gly Asp
100 105 110

gta gat tac gct cag gcg tct gcg gaa ctg gcc aaa gcg atc gcg cag 384
Val Asp Tyr Ala Gln Ala Ser Ala Glu Leu Ala Lys Ala Ile Ala Gln
115 120 125

ctg cgc gtt atc gag ttg acc aaa aaa gcg atg taa 420
Leu Arg Val Ile Glu Leu Thr Lys Lys Ala Met *
130 135

<210> 143
<211> 1383
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1383)

<400> 143
atg gct act gga aag att gtc cag gta atc ggc gcc gta gtt gac gtc 48

```
Met Ala Thr Gly Lys Ile Val Gln Val Ile Gly Ala Val Val Asp Val
 1               5                  10                  15

gaa ttc cct cag gat gcc gta ccg cgc gtg tac gat gct ctt gag gtg     96
Glu Phe Pro Gln Asp Ala Val Pro Arg Val Tyr Asp Ala Leu Glu Val
            20                  25                  30

caa aat ggt aat gag cgt ctg gtg ctg gaa gtt cag cag cag ctc ggc    144
Gln Asn Gly Asn Glu Arg Leu Val Leu Glu Val Gln Gln Gln Leu Gly
        35                  40                  45

ggc ggt atc gta cgt acc atc gca atg ggt tcc tcc gac ggt ctg cgt    192
Gly Gly Ile Val Arg Thr Ile Ala Met Gly Ser Ser Asp Gly Leu Arg
    50                  55                  60

cgc ggt ctg gat gta aaa gac ctc gaa cac ccg att gaa gtc ccg gta    240
Arg Gly Leu Asp Val Lys Asp Leu Glu His Pro Ile Glu Val Pro Val
65                  70                  75                  80

ggt aaa gcg act ctg ggc cgt atc atg aac gta ctg ggt gaa ccg gtc    288
Gly Lys Ala Thr Leu Gly Arg Ile Met Asn Val Leu Gly Glu Pro Val
                85                  90                  95

gac atg aaa ggc gag atc ggt gaa gaa gag cgt tgg gcg att cac cgc    336
Asp Met Lys Gly Glu Ile Gly Glu Glu Glu Arg Trp Ala Ile His Arg
            100                 105                 110

gca gca cct tcc tac gaa gag ctg tca aac tct cag gaa ctg ctg gaa    384
Ala Ala Pro Ser Tyr Glu Glu Leu Ser Asn Ser Gln Glu Leu Leu Glu
        115                 120                 125

acc ggt atc aaa gtt atc gac ctg atg tgt ccg ttc gct aag ggc ggt    432
Thr Gly Ile Lys Val Ile Asp Leu Met Cys Pro Phe Ala Lys Gly Gly
    130                 135                 140

aaa gtt ggt ctg ttc ggt ggt gcg ggt gta ggt aaa acc gta aac atg    480
Lys Val Gly Leu Phe Gly Gly Ala Gly Val Gly Lys Thr Val Asn Met
145                 150                 155                 160

atg gag ctc att cgt aac atc gcg atc gag cac tcc ggt tac tct gtg    528
Met Glu Leu Ile Arg Asn Ile Ala Ile Glu His Ser Gly Tyr Ser Val
                165                 170                 175

ttt gcg ggc gta ggt gaa cgt act cgt gag ggt aac gac ttc tac cac    576
Phe Ala Gly Val Gly Glu Arg Thr Arg Glu Gly Asn Asp Phe Tyr His
            180                 185                 190

gaa atg acc gac tcc aac gtt atc gac aaa gta tcc ctg gtg tat ggc    624
Glu Met Thr Asp Ser Asn Val Ile Asp Lys Val Ser Leu Val Tyr Gly
        195                 200                 205

cag atg aac gag ccg ccg gga aac cgt ctg cgc gtt gct ctg acc ggt    672
Gln Met Asn Glu Pro Pro Gly Asn Arg Leu Arg Val Ala Leu Thr Gly
    210                 215                 220

ctg acc atg gct gag aaa ttc cgt gac gaa ggt cgt gac gtt ctg ctg    720
Leu Thr Met Ala Glu Lys Phe Arg Asp Glu Gly Arg Asp Val Leu Leu
225                 230                 235                 240

ttc gtt gac aac atc tat cgt tac acc ctg gcc ggt acg gaa gta tcc    768
Phe Val Asp Asn Ile Tyr Arg Tyr Thr Leu Ala Gly Thr Glu Val Ser
```

```
                245                 250                 255

gca ctg ctg ggc cgt atg cct tca gcg gta ggt tat cag ccg acc ctg    816
Ala Leu Leu Gly Arg Met Pro Ser Ala Val Gly Tyr Gln Pro Thr Leu
            260                 265                 270

gcg gaa gag atg ggc gtt ctg cag gaa cgt atc acc tcc acc aaa act    864
Ala Glu Glu Met Gly Val Leu Gln Glu Arg Ile Thr Ser Thr Lys Thr
        275                 280                 285

ggt tct atc acc tcc gta cag gca gta tac gta cct gcg gat gac ttg    912
Gly Ser Ile Thr Ser Val Gln Ala Val Tyr Val Pro Ala Asp Asp Leu
    290                 295                 300

act gac ccg tct ccg gca acc acc ttt gcg cac ctt gac gca acc gtg    960
Thr Asp Pro Ser Pro Ala Thr Thr Phe Ala His Leu Asp Ala Thr Val
305                 310                 315                 320

gta ctg agc cgt cag atc gcg tct ctg ggt atc tac ccg gcc gtt gac    1008
Val Leu Ser Arg Gln Ile Ala Ser Leu Gly Ile Tyr Pro Ala Val Asp
                325                 330                 335

ccg ctg gac tcc acc agc cgt cag ctg gac ccg ctg gtg gtt ggt cag    1056
Pro Leu Asp Ser Thr Ser Arg Gln Leu Asp Pro Leu Val Val Gly Gln
            340                 345                 350

gaa cac tac gac acc gcg cgt ggc gtt cag tcc atc ctg caa cgt tat    1104
Glu His Tyr Asp Thr Ala Arg Gly Val Gln Ser Ile Leu Gln Arg Tyr
        355                 360                 365

cag gaa ctg aaa gac atc atc gcc atc ctg ggt atg gat gaa ctg tct    1152
Gln Glu Leu Lys Asp Ile Ile Ala Ile Leu Gly Met Asp Glu Leu Ser
    370                 375                 380

gaa gaa gac aaa ctg gtg gta gcg cgt gct cgt aag atc cag cgc ttc    1200
Glu Glu Asp Lys Leu Val Val Ala Arg Ala Arg Lys Ile Gln Arg Phe
385                 390                 395                 400

ctg tcc cag ccg ttc ttc gtg gca gaa gta ttc acc ggt tct ccg ggt    1248
Leu Ser Gln Pro Phe Phe Val Ala Glu Val Phe Thr Gly Ser Pro Gly
                405                 410                 415

aaa tac gtc tcc ctg aaa gac acc atc cgt ggc ttt aaa ggc atc atg    1296
Lys Tyr Val Ser Leu Lys Asp Thr Ile Arg Gly Phe Lys Gly Ile Met
            420                 425                 430

gaa ggc gaa tac gat cac ctg ccg gag cag gcg ttc tac atg gtc ggt    1344
Glu Gly Glu Tyr Asp His Leu Pro Glu Gln Ala Phe Tyr Met Val Gly
        435                 440                 445

tcc atc gaa gaa gct gtg gaa aaa gcc aaa aaa ctt taa                1383
Ser Ile Glu Glu Ala Val Glu Lys Ala Lys Lys Leu  *
    450                 455                 460
```

<210> 144
<211> 864
<212> DNA
<213> Escherichia coli

<220>

<221> CDS
<222> (1)...(864)

<400> 144
atg gcc ggc gca aaa gag ata cgt agt aag atc gca agc gtc cag aac 48
Met Ala Gly Ala Lys Glu Ile Arg Ser Lys Ile Ala Ser Val Gln Asn
1 5 10 15

acg caa aag atc act aaa gcg atg gag atg gtc gcc gct tcc aaa atg 96
Thr Gln Lys Ile Thr Lys Ala Met Glu Met Val Ala Ala Ser Lys Met
20 25 30

cgt aaa tcg cag gat cgc atg gcg gcc agc cgt cct tat gca gaa acc 144
Arg Lys Ser Gln Asp Arg Met Ala Ala Ser Arg Pro Tyr Ala Glu Thr
35 40 45

atg cgc aaa gtg att ggt cac ctt gca cac ggt aat ctg gaa tat aag 192
Met Arg Lys Val Ile Gly His Leu Ala His Gly Asn Leu Glu Tyr Lys
50 55 60

cac cct tac ctg gaa gac cgc gac gtt aaa cgc gtg ggc tac ctg gtg 240
His Pro Tyr Leu Glu Asp Arg Asp Val Lys Arg Val Gly Tyr Leu Val
65 70 75 80

gtg tcg acc gac cgt ggt ttg tgc ggt ggt ttg aac att aac ctg ttc 288
Val Ser Thr Asp Arg Gly Leu Cys Gly Gly Leu Asn Ile Asn Leu Phe
85 90 95

aaa aaa ctg ctg gcg gaa atg aag acc tgg acc gac aaa ggc gtt caa 336
Lys Lys Leu Leu Ala Glu Met Lys Thr Trp Thr Asp Lys Gly Val Gln
100 105 110

tgc gac ctc gca atg atc ggc tcg aaa ggc gtg tcg ttc ttc aac tcc 384
Cys Asp Leu Ala Met Ile Gly Ser Lys Gly Val Ser Phe Phe Asn Ser
115 120 125

gtg ggc ggc aat gtt gtt gcc cag gtc acc ggc atg ggg gat aac cct 432
Val Gly Gly Asn Val Val Ala Gln Val Thr Gly Met Gly Asp Asn Pro
130 135 140

tcc ctg tcc gaa ctg atc ggt ccg gta aaa gtg atg ttg cag gcc tac 480
Ser Leu Ser Glu Leu Ile Gly Pro Val Lys Val Met Leu Gln Ala Tyr
145 150 155 160

gac gaa ggc cgt ctg gac aag ctt tac att gtc agc aac aaa ttt att 528
Asp Glu Gly Arg Leu Asp Lys Leu Tyr Ile Val Ser Asn Lys Phe Ile
165 170 175

aac acc atg tct cag gtt ccg acc atc agc cag ctg ctg ccg tta ccg 576
Asn Thr Met Ser Gln Val Pro Thr Ile Ser Gln Leu Leu Pro Leu Pro
180 185 190

gca tca gat gat gat gat ctg aaa cat aaa tcc tgg gat tac ctg tac 624
Ala Ser Asp Asp Asp Asp Leu Lys His Lys Ser Trp Asp Tyr Leu Tyr
195 200 205

gaa ccc gat ccg aag gcg ttg ctg gat acc ctg ctg cgt cgt tat gtc 672
Glu Pro Asp Pro Lys Ala Leu Leu Asp Thr Leu Leu Arg Arg Tyr Val
210 215 220

gaa tct cag gtt tat cag ggc gtg gtt gaa aac ctg gcc agc gag cag 720

```
Glu Ser Gln Val Tyr Gln Gly Val Val Glu Asn Leu Ala Ser Glu Gln
225                 230                 235                 240

gcc gcc cgt atg gtg gcg atg aaa gcc gcg acc gac aat ggc ggc agc    768
Ala Ala Arg Met Val Ala Met Lys Ala Ala Thr Asp Asn Gly Gly Ser
                245                 250                 255

ctg att aaa gag ctg cag ttg gta tac aac aaa gct cgt cag gcc agc    816
Leu Ile Lys Glu Leu Gln Leu Val Tyr Asn Lys Ala Arg Gln Ala Ser
            260                 265                 270

att act cag gaa ctc acc gag atc gtc tcg ggg gcc gcc gcg gtt taa    864
Ile Thr Gln Glu Leu Thr Glu Ile Val Ser Gly Ala Ala Ala Val  *
        275                 280                 285
```

<210> 145
<211> 1542
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1542)

<400> 145

```
atg caa ctg aat tcc acc gaa atc agc gaa ctg atc aag cag cgc att    48
Met Gln Leu Asn Ser Thr Glu Ile Ser Glu Leu Ile Lys Gln Arg Ile
 1               5                  10                  15

gct cag ttc aat gtt gtg agt gaa gct cac aac gaa ggt act att gtt    96
Ala Gln Phe Asn Val Val Ser Glu Ala His Asn Glu Gly Thr Ile Val
            20                  25                  30

tct gta agt gac ggt gtt atc cgc att cac ggc ctg gcc gat tgt atg    144
Ser Val Ser Asp Gly Val Ile Arg Ile His Gly Leu Ala Asp Cys Met
        35                  40                  45

cag ggt gaa atg atc tcc ctg ccg ggt aac cgt tac gct atc gca ctg    192
Gln Gly Glu Met Ile Ser Leu Pro Gly Asn Arg Tyr Ala Ile Ala Leu
    50                  55                  60

aac ctc gag cgc gac tct gta ggt gcg gtt gtt atg ggt ccg tac gct    240
Asn Leu Glu Arg Asp Ser Val Gly Ala Val Val Met Gly Pro Tyr Ala
 65                  70                  75                  80

gac ctt gcc gaa ggc atg aaa gtt aag tgt act ggc cgt atc ctg gaa    288
Asp Leu Ala Glu Gly Met Lys Val Lys Cys Thr Gly Arg Ile Leu Glu
                85                  90                  95

gtt ccg gtt ggc cgt ggc ctg ctg ggc cgt gtg gtt aac act ctg ggt    336
Val Pro Val Gly Arg Gly Leu Leu Gly Arg Val Val Asn Thr Leu Gly
            100                 105                 110

gca cca atc gac ggt aaa ggt ccg ctg gat cac gac ggc ttc tct gct    384
Ala Pro Ile Asp Gly Lys Gly Pro Leu Asp His Asp Gly Phe Ser Ala
        115                 120                 125

gta gaa gca atc gct ccg ggc gtt atc gaa cgt cag tcc gta gat cag    432
Val Glu Ala Ile Ala Pro Gly Val Ile Glu Arg Gln Ser Val Asp Gln
```

```
    130                 135                 140

ccg gta cag acc ggt tat aaa gcc gtt gac tcc atg atc cca atc ggt      480
Pro Val Gln Thr Gly Tyr Lys Ala Val Asp Ser Met Ile Pro Ile Gly
145                 150                 155                 160

cgt ggt cag cgt gaa ttg atc atc ggt gac cgt cag aca ggt aaa acc      528
Arg Gly Gln Arg Glu Leu Ile Ile Gly Asp Arg Gln Thr Gly Lys Thr
                165                 170                 175

gca ctg gct atc gat gcc atc atc aac cag cgc gat tcc ggt atc aaa      576
Ala Leu Ala Ile Asp Ala Ile Ile Asn Gln Arg Asp Ser Gly Ile Lys
            180                 185                 190

tgt atc tat gtc gct atc ggc cag aaa gcg tcc acc att tct aac gtg      624
Cys Ile Tyr Val Ala Ile Gly Gln Lys Ala Ser Thr Ile Ser Asn Val
        195                 200                 205

gta cgt aaa ctg gaa gag cac ggc gca ctg gct aac acc atc gtt gtg      672
Val Arg Lys Leu Glu Glu His Gly Ala Leu Ala Asn Thr Ile Val Val
    210                 215                 220

gta gca acc gcg tct gaa tcc gct gca ctg caa tac ctg gca ccg tat      720
Val Ala Thr Ala Ser Glu Ser Ala Ala Leu Gln Tyr Leu Ala Pro Tyr
225                 230                 235                 240

gcc ggt tgc gca atg ggc gaa tac ttc cgt gac cgc ggt gaa gat gcg      768
Ala Gly Cys Ala Met Gly Glu Tyr Phe Arg Asp Arg Gly Glu Asp Ala
                245                 250                 255

ctg atc att tac gat gac ctg tct aaa cag gct gtt gct tac cgt cag      816
Leu Ile Ile Tyr Asp Asp Leu Ser Lys Gln Ala Val Ala Tyr Arg Gln
            260                 265                 270

atc tcc ctg ctg ctc cgt cgt ccg cca gga cgt gaa gca ttc ccg ggc      864
Ile Ser Leu Leu Leu Arg Arg Pro Pro Gly Arg Glu Ala Phe Pro Gly
        275                 280                 285

gac gtt ttc tac ctc cac tct cgt ctg ctg gag cgt gct gca cgt gtt      912
Asp Val Phe Tyr Leu His Ser Arg Leu Leu Glu Arg Ala Ala Arg Val
    290                 295                 300

aac gcc gaa tac gtt gaa gcc ttc acc aaa ggt gaa gtg aaa ggg aaa      960
Asn Ala Glu Tyr Val Glu Ala Phe Thr Lys Gly Glu Val Lys Gly Lys
305                 310                 315                 320

acc ggt tct ctg acc gca ctg ccg att atc gaa act cag gcg ggt gac      1008
Thr Gly Ser Leu Thr Ala Leu Pro Ile Ile Glu Thr Gln Ala Gly Asp
                325                 330                 335

gtt tct gcg ttc gtt ccg acc aac gta atc tcc att acc gat ggt cag      1056
Val Ser Ala Phe Val Pro Thr Asn Val Ile Ser Ile Thr Asp Gly Gln
            340                 345                 350

atc ttc ctg gaa acc aac ctg ttc aac gcc ggt att cgt cct gcg gtt      1104
Ile Phe Leu Glu Thr Asn Leu Phe Asn Ala Gly Ile Arg Pro Ala Val
        355                 360                 365

aac ccg ggt att tcc gta tcc cgt gtt ggt ggt gca gca cag acc aag      1152
Asn Pro Gly Ile Ser Val Ser Arg Val Gly Gly Ala Ala Gln Thr Lys
    370                 375                 380
```

```
atc atg aaa aaa ctg tcc ggt ggt atc cgt acc gct ctg gca cag tat    1200
Ile Met Lys Lys Leu Ser Gly Gly Ile Arg Thr Ala Leu Ala Gln Tyr
385                 390                 395                 400

cgt gaa ctg gca gcg ttc tct cag ttt gca tcc gac ctt gac gat gca    1248
Arg Glu Leu Ala Ala Phe Ser Gln Phe Ala Ser Asp Leu Asp Asp Ala
                405                 410                 415

aca cgt aag cag ctt gac cac ggt cag aaa gtg acc gaa ctg ctg aaa    1296
Thr Arg Lys Gln Leu Asp His Gly Gln Lys Val Thr Glu Leu Leu Lys
            420                 425                 430

cag aaa cag tat gcg ccg atg tcc gtt gcg cag cag tct ctg gtt ctg    1344
Gln Lys Gln Tyr Ala Pro Met Ser Val Ala Gln Gln Ser Leu Val Leu
        435                 440                 445

ttc gca gca gaa cgt ggt tac ctg gcg gat gtt gaa ctg tcg aaa att    1392
Phe Ala Ala Glu Arg Gly Tyr Leu Ala Asp Val Glu Leu Ser Lys Ile
    450                 455                 460

ggc agc ttc gaa gcc gct ctg ctg gct tac gtc gac cgt gat cac gct    1440
Gly Ser Phe Glu Ala Ala Leu Leu Ala Tyr Val Asp Arg Asp His Ala
465                 470                 475                 480

ccg ttg atg caa gag atc aac cag acc ggt ggc tac aac gac gaa atc    1488
Pro Leu Met Gln Glu Ile Asn Gln Thr Gly Gly Tyr Asn Asp Glu Ile
                485                 490                 495

gaa ggc aag ctg aaa ggc atc ctc gat tcc ttc aaa gca acc caa tcc    1536
Glu Gly Lys Leu Lys Gly Ile Leu Asp Ser Phe Lys Ala Thr Gln Ser
            500                 505                 510

tgg taa                                                            1542
Trp  *
```

<210> 146
<211> 534
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(534)

<400> 146

```
atg tct gaa ttt att acg gta gct cgc ccc tac gcc aaa gca gct ttt    48
Met Ser Glu Phe Ile Thr Val Ala Arg Pro Tyr Ala Lys Ala Ala Phe
 1               5                  10                  15

gac ttt gcc gtc gaa cac caa agt gta gaa cgc tgg cag gac atg ctg    96
Asp Phe Ala Val Glu His Gln Ser Val Glu Arg Trp Gln Asp Met Leu
            20                  25                  30

gcg ttt gcc gcc gag gta acc aaa aac gaa caa atg gca gag ctt ctc    144
Ala Phe Ala Ala Glu Val Thr Lys Asn Glu Gln Met Ala Glu Leu Leu
        35                  40                  45

tct ggc gcg ctt gcg cca gaa acg ctc gcc gag tcg ttt atc gca gtt    192
```

```
Ser Gly Ala Leu Ala Pro Glu Thr Leu Ala Glu Ser Phe Ile Ala Val
     50                  55                  60

tgt ggt gag caa ctg gac gaa aac ggt cag aac ctg att cgg gtt atg     240
Cys Gly Glu Gln Leu Asp Glu Asn Gly Gln Asn Leu Ile Arg Val Met
 65                  70                  75                  80

gct gaa aat ggt cgt ctt aac gcg ctc ccg gat gtt ctg gag cag ttt     288
Ala Glu Asn Gly Arg Leu Asn Ala Leu Pro Asp Val Leu Glu Gln Phe
                 85                  90                  95

att cac ctg cgt gcc gtg agt gag gct acc gct gag gta gac gtc att     336
Ile His Leu Arg Ala Val Ser Glu Ala Thr Ala Glu Val Asp Val Ile
            100                 105                 110

tcc gct gcc gca ctg agt gaa caa cag ctc gcg aaa att tct gct gcg     384
Ser Ala Ala Ala Leu Ser Glu Gln Gln Leu Ala Lys Ile Ser Ala Ala
        115                 120                 125

atg gaa aaa cgt ctg tca cgc aaa gtt aag ctg aat tgc aaa atc gat     432
Met Glu Lys Arg Leu Ser Arg Lys Val Lys Leu Asn Cys Lys Ile Asp
    130                 135                 140

aag tct gta atg gca ggc gtt atc atc cga gcg ggt gat atg gtc att     480
Lys Ser Val Met Ala Gly Val Ile Ile Arg Ala Gly Asp Met Val Ile
145                 150                 155                 160

gat ggc agc gta cgc ggt cgt ctt gag cgc ctt gca gac gtc ttg cag     528
Asp Gly Ser Val Arg Gly Arg Leu Glu Arg Leu Ala Asp Val Leu Gln
                165                 170                 175

tct taa                                                             534
Ser  *
```

<210> 147
<211> 471
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(471)

<400> 147

```
gtg aat ctt aac gca aca atc ctc ggc cag gcc atc gcg ttt gtc ctg     48
Met Asn Leu Asn Ala Thr Ile Leu Gly Gln Ala Ile Ala Phe Val Leu
 1               5                  10                  15

ttc gtt ctg ttc tgc atg aag tac gta tgg ccg cca tta atg gca gcc     96
Phe Val Leu Phe Cys Met Lys Tyr Val Trp Pro Pro Leu Met Ala Ala
            20                  25                  30

atc gaa aaa cgt caa aaa gaa att gct gac ggc ctt gct tcc gca gaa     144
Ile Glu Lys Arg Gln Lys Glu Ile Ala Asp Gly Leu Ala Ser Ala Glu
        35                  40                  45

cga gca cat aag gac ctt gac ctt gca aag gcc agc gcg acc gac cag     192
Arg Ala His Lys Asp Leu Asp Leu Ala Lys Ala Ser Ala Thr Asp Gln
    50                  55                  60
```

```
ctg aaa aaa gcg aaa gcg gaa gcc cag gta atc atc gag cag gcg aac    240
Leu Lys Lys Ala Lys Ala Glu Ala Gln Val Ile Ile Glu Gln Ala Asn
 65                  70                  75                  80

aaa cgc cgc tcg cag att ctg gac gaa gcg aaa gct gag gca gaa cag    288
Lys Arg Arg Ser Gln Ile Leu Asp Glu Ala Lys Ala Glu Ala Glu Gln
                 85                  90                  95

gaa cgt act aaa atc gtg gcc cag gcg cag gcg gaa att gaa gcc gag    336
Glu Arg Thr Lys Ile Val Ala Gln Ala Gln Ala Glu Ile Glu Ala Glu
            100                 105                 110

cgt aaa cgt gcc cgt gaa gag ctg cgt aag caa gtt gct atc ctg gct    384
Arg Lys Arg Ala Arg Glu Glu Leu Arg Lys Gln Val Ala Ile Leu Ala
        115                 120                 125

gtt gct ggc gcc gag aag atc atc gaa cgt tcc gtg gat gaa gct gct    432
Val Ala Gly Ala Glu Lys Ile Ile Glu Arg Ser Val Asp Glu Ala Ala
    130                 135                 140

aac agc gac atc gtg gat aaa ctt gtc gct gaa ctg taa                471
Asn Ser Asp Ile Val Asp Lys Leu Val Ala Glu Leu  *
145                 150                 155
```

<210> 148
<211> 240
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(240)

<400> 148

```
atg gaa aac ctg aat atg gat ctg ctg tac atg gct gcc gct gtg atg    48
Met Glu Asn Leu Asn Met Asp Leu Leu Tyr Met Ala Ala Ala Val Met
 1               5                  10                  15

atg ggt ctg gcg gca atc ggt gct gcg atc ggt atc ggc atc ctc ggg    96
Met Gly Leu Ala Ala Ile Gly Ala Ala Ile Gly Ile Gly Ile Leu Gly
            20                  25                  30

ggt aaa ttc ctg gaa ggc gca gcg cgt caa cct gat ctg att cct ctg    144
Gly Lys Phe Leu Glu Gly Ala Ala Arg Gln Pro Asp Leu Ile Pro Leu
        35                  40                  45

ctg cgt act cag ttc ttt atc gtt atg ggt ctg gtg gat gct atc ccg    192
Leu Arg Thr Gln Phe Phe Ile Val Met Gly Leu Val Asp Ala Ile Pro
    50                  55                  60

atg atc gct gta ggt ctg ggt ctg tac gtg atg ttc gct gtc gcg tag    240
Met Ile Ala Val Gly Leu Gly Leu Tyr Val Met Phe Ala Val Ala  *
 65                  70                  75
```

<210> 149
<211> 816
<212> DNA

<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(816)

<400> 149
atg gct tca gaa aat atg acg ccg cag gat tac ata gga cac cac ctg 48
Met Ala Ser Glu Asn Met Thr Pro Gln Asp Tyr Ile Gly His His Leu
1 5 10 15

aat aac ctt cag ctg gac ctg cgt aca ttc tcg ctg gtg gat cca caa 96
Asn Asn Leu Gln Leu Asp Leu Arg Thr Phe Ser Leu Val Asp Pro Gln
20 25 30

aac ccc cca gcc acc ttc tgg aca atc aat att gac tcc atg ttc ttc 144
Asn Pro Pro Ala Thr Phe Trp Thr Ile Asn Ile Asp Ser Met Phe Phe
35 40 45

tcg gtg gtg ctg ggt ctg ttg ttc ctg gtt tta ttc cgt agc gta gcc 192
Ser Val Val Leu Gly Leu Leu Phe Leu Val Leu Phe Arg Ser Val Ala
50 55 60

aaa aag gcg acc agc ggt gtg cca ggt aag ttt cag acc gcg att gag 240
Lys Lys Ala Thr Ser Gly Val Pro Gly Lys Phe Gln Thr Ala Ile Glu
65 70 75 80

ctg gtg atc ggc ttt gtt aat ggt agc gtg aaa gac atg tac cat ggc 288
Leu Val Ile Gly Phe Val Asn Gly Ser Val Lys Asp Met Tyr His Gly
85 90 95

aaa agc aag ctg att gct ccg ctg gcc ctg acg atc ttc gtc tgg gta 336
Lys Ser Lys Leu Ile Ala Pro Leu Ala Leu Thr Ile Phe Val Trp Val
100 105 110

ttc ctg atg aac ctg atg gat tta ctg cct atc gac ctg ctg ccg tac 384
Phe Leu Met Asn Leu Met Asp Leu Leu Pro Ile Asp Leu Leu Pro Tyr
115 120 125

att gct gaa cat gta ctg ggt ctg cct gca ctg cgt gtg gtt ccg tct 432
Ile Ala Glu His Val Leu Gly Leu Pro Ala Leu Arg Val Val Pro Ser
130 135 140

gcg gac gtg aac gta acg ctg tct atg gca ctg ggc gta ttt atc ctg 480
Ala Asp Val Asn Val Thr Leu Ser Met Ala Leu Gly Val Phe Ile Leu
145 150 155 160

att ctg ttc tac agc atc aaa atg aaa ggc atc ggc ggc ttc acg aaa 528
Ile Leu Phe Tyr Ser Ile Lys Met Lys Gly Ile Gly Gly Phe Thr Lys
165 170 175

gag ttg acg ctg cag ccg ttc aat cac tgg gcg ttc att cct gtc aac 576
Glu Leu Thr Leu Gln Pro Phe Asn His Trp Ala Phe Ile Pro Val Asn
180 185 190

tta atc ctt gaa ggg gta agc ctg ctg tcc aaa cca gtt tca ctc ggt 624
Leu Ile Leu Glu Gly Val Ser Leu Leu Ser Lys Pro Val Ser Leu Gly
195 200 205

ttg cga ctg ttc ggt aac atg tat gcc ggt gag ctg att ttc att ctg 672
Leu Arg Leu Phe Gly Asn Met Tyr Ala Gly Glu Leu Ile Phe Ile Leu

```
    210                 215                 220

att gct ggt ctg ttg ccg tgg tgg tca cag tgg atc ctg aat gtg ccg    720
Ile Ala Gly Leu Leu Pro Trp Trp Ser Gln Trp Ile Leu Asn Val Pro
225                 230                 235                 240

tgg gcc att ttc cac atc ctg atc att acg ctg caa gcc ttc atc ttc    768
Trp Ala Ile Phe His Ile Leu Ile Ile Thr Leu Gln Ala Phe Ile Phe
                245                 250                 255

atg gtt ctg acg atc gtc tat ctg tcg atg gcg tct gaa gaa cat taa    816
Met Val Leu Thr Ile Val Tyr Leu Ser Met Ala Ser Glu Glu His  *
            260                 265                 270
```

<210> 150
<211> 393
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(393)

<400> 150

```
gtg aaa aac gtg atg tct gtg tcg ctc gtg agt cga aac gtt gct cgg    48
Met Lys Asn Val Met Ser Val Ser Leu Val Ser Arg Asn Val Ala Arg
 1               5                  10                  15

aag ctt ctg ctc gtt cag tta ctg gtg gtg ata gca agt gga ttg ctg    96
Lys Leu Leu Leu Val Gln Leu Leu Val Val Ile Ala Ser Gly Leu Leu
            20                  25                  30

ttc agc ctc aaa gac ccc ttc tgg ggc gtc tct gca ata agc ggg ggc    144
Phe Ser Leu Lys Asp Pro Phe Trp Gly Val Ser Ala Ile Ser Gly Gly
        35                  40                  45

ctg gca gtc ttt ctg cct aac gtt ttg ttt atg ata ttt gcc tgg cgt    192
Leu Ala Val Phe Leu Pro Asn Val Leu Phe Met Ile Phe Ala Trp Arg
    50                  55                  60

cac cag gcg cat aca cca gcg aaa ggc cgg gtg gcc tgg aca ttc gca    240
His Gln Ala His Thr Pro Ala Lys Gly Arg Val Ala Trp Thr Phe Ala
 65                 70                  75                  80

ttt ggc gaa gct ttc aaa gtt ctg gcg atg ttg gtg tta ctg gtg gtg    288
Phe Gly Glu Ala Phe Lys Val Leu Ala Met Leu Val Leu Leu Val Val
                85                  90                  95

gcg ttg gcg gtt tta aag gcg gta ttc ttg ccg ctg atc gtt acg tgg    336
Ala Leu Ala Val Leu Lys Ala Val Phe Leu Pro Leu Ile Val Thr Trp
            100                 105                 110

gtt ttg gtg ctg gtg gtt cag ata ctg gca ccg gct gta att aac aac    384
Val Leu Val Leu Val Val Gln Ile Leu Ala Pro Ala Val Ile Asn Asn
        115                 120                 125

aaa ggg taa                                                        393
Lys Gly  *
    130
```

<210> 151
<211> 1242
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1242)

<400> 151
atg caa aaa cac ggc gac cgc tat gtg tgg atc aac ccg cct gct att 48
Met Gln Lys His Gly Asp Arg Tyr Val Trp Ile Asn Pro Pro Ala Ile
1 5 10 15

ccg ctt tct acc gaa gag atg gac agc gtt ttt gcg ctg cca tac aag 96
Pro Leu Ser Thr Glu Glu Met Asp Ser Val Phe Ala Leu Pro Tyr Lys
20 25 30

cgc gtg cca cat ccg gcc tat ggc aat gcc cgt att ccg gct tac gaa 144
Arg Val Pro His Pro Ala Tyr Gly Asn Ala Arg Ile Pro Ala Tyr Glu
35 40 45

atg atc cgt ttt tcg gtc aac att atg cgt ggc tgc ttt ggc ggc tgc 192
Met Ile Arg Phe Ser Val Asn Ile Met Arg Gly Cys Phe Gly Gly Cys
50 55 60

tct ttc tgt tct atc acc gag cac gaa ggg cgc att att cag agc cgt 240
Ser Phe Cys Ser Ile Thr Glu His Glu Gly Arg Ile Ile Gln Ser Arg
65 70 75 80

tcc gaa gat tcg atc att aat gag atc gaa gcg atc cgc gac acc gtt 288
Ser Glu Asp Ser Ile Ile Asn Glu Ile Glu Ala Ile Arg Asp Thr Val
85 90 95

cca ggt ttt acg ggc gtg att tcc gat ctt ggt ggg cca act gcc aac 336
Pro Gly Phe Thr Gly Val Ile Ser Asp Leu Gly Gly Pro Thr Ala Asn
100 105 110

atg tat atg ttg cgc tgc aaa tcg cca cgc gct gaa caa act tgt cgc 384
Met Tyr Met Leu Arg Cys Lys Ser Pro Arg Ala Glu Gln Thr Cys Arg
115 120 125

cgt ttg tcg tgc gtt tat ccg gat att tgt ccg cac atg gac acg aac 432
Arg Leu Ser Cys Val Tyr Pro Asp Ile Cys Pro His Met Asp Thr Asn
130 135 140

cac gaa cca acg atc aac ctc tat cgc cgt gcg cgt gat ctg aaa ggc 480
His Glu Pro Thr Ile Asn Leu Tyr Arg Arg Ala Arg Asp Leu Lys Gly
145 150 155 160

att aaa aag atc ctg att gcc tct ggt gtg cgt tat gac ata gcc gta 528
Ile Lys Lys Ile Leu Ile Ala Ser Gly Val Arg Tyr Asp Ile Ala Val
165 170 175

gaa gat ccg cgc tat atc aaa gaa ctg gcg acc cat cac gtc ggc ggt 576
Glu Asp Pro Arg Tyr Ile Lys Glu Leu Ala Thr His His Val Gly Gly
180 185 190

tat ctg aag att gcc ccg gaa cat acc gaa gaa ggg ccg tta tcg aag 624

```
Tyr Leu Lys Ile Ala Pro Glu His Thr Glu Glu Gly Pro Leu Ser Lys
        195                 200                 205

atg atg aag ccg ggc atg ggc agc tat gac cgc ttt aaa gag ctg ttc      672
Met Met Lys Pro Gly Met Gly Ser Tyr Asp Arg Phe Lys Glu Leu Phe
    210                 215                 220

gat act tac tca aaa cag gca ggt aaa gag cag tat ctg atc ccc tat      720
Asp Thr Tyr Ser Lys Gln Ala Gly Lys Glu Gln Tyr Leu Ile Pro Tyr
225                 230                 235                 240

ttc atc tcc gcg cac ccc ggt acg cgt gat gaa gat atg gtg aat ctg      768
Phe Ile Ser Ala His Pro Gly Thr Arg Asp Glu Asp Met Val Asn Leu
                245                 250                 255

gcg ctg tgg ctg aaa aag cac cgc ttc cgc ctc gac cag gtg cag aac      816
Ala Leu Trp Leu Lys Lys His Arg Phe Arg Leu Asp Gln Val Gln Asn
            260                 265                 270

ttc tat ccg tcg ccg ctg gcg aac tca acc acc atg tat tac acc ggg      864
Phe Tyr Pro Ser Pro Leu Ala Asn Ser Thr Thr Met Tyr Tyr Thr Gly
        275                 280                 285

aaa aac ccg ctg gcg aag att ggt tat aag agc gaa gac gtc ttc gta      912
Lys Asn Pro Leu Ala Lys Ile Gly Tyr Lys Ser Glu Asp Val Phe Val
    290                 295                 300

ccg aag ggc gac aaa cag cgt cgt ttg cat aaa gcg ttg ttg cgt tac      960
Pro Lys Gly Asp Lys Gln Arg Arg Leu His Lys Ala Leu Leu Arg Tyr
305                 310                 315                 320

cac gat ccg gca aac tgg ccg tta atc cgc cag gcg ctg gaa gcg atg      1008
His Asp Pro Ala Asn Trp Pro Leu Ile Arg Gln Ala Leu Glu Ala Met
                325                 330                 335

ggc aaa aag cat ctg att ggc agc cgt cgc gat tgc tta gtg cct gcg      1056
Gly Lys Lys His Leu Ile Gly Ser Arg Arg Asp Cys Leu Val Pro Ala
            340                 345                 350

cca acc att gaa gag atg cgt gaa gct cgt cgc cag aac cgc aat acc      1104
Pro Thr Ile Glu Glu Met Arg Glu Ala Arg Arg Gln Asn Arg Asn Thr
        355                 360                 365

cgt ccg gcg ttg acg aaa cat acg ccg atg gcg acc cag cgt cag acg      1152
Arg Pro Ala Leu Thr Lys His Thr Pro Met Ala Thr Gln Arg Gln Thr
    370                 375                 380

cct gct acg gca aaa aaa gcg tcg tct acg caa tct cgt ccg gtg aat      1200
Pro Ala Thr Ala Lys Lys Ala Ser Ser Thr Gln Ser Arg Pro Val Asn
385                 390                 395                 400

gct ggt gcg aag aaa cgg cct aaa gcg gcg gtt gga cgt taa              1242
Ala Gly Ala Lys Lys Arg Pro Lys Ala Ala Val Gly Arg  *
                405                 410
```

<210> 152
<211> 978
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(978)

<400> 152
atg agc tct atc tcc ctg atc caa ccg gat cgc gac ctg ttc tcc tgg 48
Met Ser Ser Ile Ser Leu Ile Gln Pro Asp Arg Asp Leu Phe Ser Trp
1 5 10 15

ccg cag tac tgg gcc gcc tgt ttt gga ccg gca ccg ttt ttg ccg atg 96
Pro Gln Tyr Trp Ala Ala Cys Phe Gly Pro Ala Pro Phe Leu Pro Met
20 25 30

tct cgt gaa gag atg gat caa ctt ggc tgg gat agc tgc gac atc att 144
Ser Arg Glu Glu Met Asp Gln Leu Gly Trp Asp Ser Cys Asp Ile Ile
35 40 45

ttg gtt act ggc gac gcg tat gtc gat cac cca agc ttc ggg atg gcg 192
Leu Val Thr Gly Asp Ala Tyr Val Asp His Pro Ser Phe Gly Met Ala
50 55 60

att tgc ggt cgt atg ctg gaa gcg cag ggc ttt cgc gtc ggg atc atc 240
Ile Cys Gly Arg Met Leu Glu Ala Gln Gly Phe Arg Val Gly Ile Ile
65 70 75 80

gcc cag ccg gac tgg agc agc aaa gac gac ttt atg cgt ctg ggt aaa 288
Ala Gln Pro Asp Trp Ser Ser Lys Asp Asp Phe Met Arg Leu Gly Lys
85 90 95

ccg aat ctg ttt ttc ggt gtt act gct ggc aac atg gat tcg atg atc 336
Pro Asn Leu Phe Phe Gly Val Thr Ala Gly Asn Met Asp Ser Met Ile
100 105 110

aac cgt tat acc gcc gat cgc cgt tta cgt cat gac gat gcc tac acg 384
Asn Arg Tyr Thr Ala Asp Arg Arg Leu Arg His Asp Asp Ala Tyr Thr
115 120 125

ccg gat aac gtc gcg ggt aag cgc ccg gat cgc gcc aca ctg gtt tat 432
Pro Asp Asn Val Ala Gly Lys Arg Pro Asp Arg Ala Thr Leu Val Tyr
130 135 140

acc cag cgt tgt aaa gag gcg tgg aaa gat gta ccg gtg atc ctc ggc 480
Thr Gln Arg Cys Lys Glu Ala Trp Lys Asp Val Pro Val Ile Leu Gly
145 150 155 160

ggt att gag gct agt ctg cgc cgt acc gcg cat tat gat tac tgg tcc 528
Gly Ile Glu Ala Ser Leu Arg Arg Thr Ala His Tyr Asp Tyr Trp Ser
165 170 175

gat acc gtg cgc cgt tcc gtg ctg gtg gat tcg aaa gcc gac atg ctg 576
Asp Thr Val Arg Arg Ser Val Leu Val Asp Ser Lys Ala Asp Met Leu
180 185 190

atg ttt ggt aac ggt gag cgt ccg ctg gtg gaa gtg gcg cat cgt ctg 624
Met Phe Gly Asn Gly Glu Arg Pro Leu Val Glu Val Ala His Arg Leu
195 200 205

gcg atg ggc gag cca att agt gaa atc cgc gat gtg cgt aat acc gcg 672
Ala Met Gly Glu Pro Ile Ser Glu Ile Arg Asp Val Arg Asn Thr Ala
210 215 220

```
att atc gtg aaa gag gcg ctg cct ggc tgg agc ggc gtg gat tcc acc    720
Ile Ile Val Lys Glu Ala Leu Pro Gly Trp Ser Gly Val Asp Ser Thr
225                 230                 235                 240

cgt ctt gat acc cct gga aaa atc gac cca atc ccg cat ccg tat ggt    768
Arg Leu Asp Thr Pro Gly Lys Ile Asp Pro Ile Pro His Pro Tyr Gly
                245                 250                 255

gaa gat ttg ccg tgc gcg gat aac aaa ccg gtg gca ccg aaa aag cag    816
Glu Asp Leu Pro Cys Ala Asp Asn Lys Pro Val Ala Pro Lys Lys Gln
            260                 265                 270

gaa gcc aaa gcc gta acc gtg cag cca ccg cgc ccg aaa ccg tgg gaa    864
Glu Ala Lys Ala Val Thr Val Gln Pro Pro Arg Pro Lys Pro Trp Glu
        275                 280                 285

aaa acc tac gtg ttg ctg cct tct ttc gag aaa gtg aag ggc gat aaa    912
Lys Thr Tyr Val Leu Leu Pro Ser Phe Glu Lys Val Lys Gly Asp Lys
    290                 295                 300

gtg ctg tac gcc cat gct tcg cgt att ctg cac cac gaa acc aac cca    960
Val Leu Tyr Ala His Ala Ser Arg Ile Leu His His Glu Thr Asn Pro
305                 310                 315                 320

gct gtg ccc gcg cat tga                                            978
Ala Val Pro Ala His  *
                325
```

<210> 153
<211> 681
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(681)

<400> 153

```
atg att cag tat ctg aac gtc ttt ttt tac gat atc tac ccg tac att    48
Met Ile Gln Tyr Leu Asn Val Phe Phe Tyr Asp Ile Tyr Pro Tyr Ile
 1               5                  10                  15

tgt gcg acg gtg ttt ttc ctc ggc agc tgg ctg cgc tac gac tac ggg    96
Cys Ala Thr Val Phe Phe Leu Gly Ser Trp Leu Arg Tyr Asp Tyr Gly
            20                  25                  30

cag tac acc tgg cgc gcc tcc tca agt caa atg ctc gat aaa cgc ggg    144
Gln Tyr Thr Trp Arg Ala Ser Ser Ser Gln Met Leu Asp Lys Arg Gly
        35                  40                  45

atg gtg ata tgg tcg aat ctg ttc cat atc ggc att ttg ggg att ttc    192
Met Val Ile Trp Ser Asn Leu Phe His Ile Gly Ile Leu Gly Ile Phe
    50                  55                  60

ttc ggg cac ttg ttc ggc atg tta acg ccg cac tgg atg tac gcg tgg    240
Phe Gly His Leu Phe Gly Met Leu Thr Pro His Trp Met Tyr Ala Trp
65                  70                  75                  80

ttt ttg cca gtg gca gcg aaa caa ctc atg gcg atg gtg ctc ggc ggt    288
```

```
Phe Leu Pro Val Ala Ala Lys Gln Leu Met Ala Met Val Leu Gly Gly
                85                  90                  95

att tgc ggt gtg ttg acg ttg att ggc ggc gca gga ctg ctg tgg cgc     336
Ile Cys Gly Val Leu Thr Leu Ile Gly Gly Ala Gly Leu Leu Trp Arg
            100                 105                 110

agg ctg aca aat cag cgc gtg cgt gcc act tcc act acg ccg gat atc     384
Arg Leu Thr Asn Gln Arg Val Arg Ala Thr Ser Thr Thr Pro Asp Ile
        115                 120                 125

atc atc atg agc att ttg ctg atc cag tgt ctg ttg ggg cta agt acc     432
Ile Ile Met Ser Ile Leu Leu Ile Gln Cys Leu Leu Gly Leu Ser Thr
    130                 135                 140

ata ccg ttt tca gca cag tat cct gat ggt agc gaa atg atg aag ctg     480
Ile Pro Phe Ser Ala Gln Tyr Pro Asp Gly Ser Glu Met Met Lys Leu
145                 150                 155                 160

gtg ggt tgg gcg caa agc att gtc act ttc cgt ggt gga tcg tca gag     528
Val Gly Trp Ala Gln Ser Ile Val Thr Phe Arg Gly Gly Ser Ser Glu
                165                 170                 175

atg ctc aac ggc gta gcg ttc gtc ttc cgc ctg cat ctg gtg ttg gga     576
Met Leu Asn Gly Val Ala Phe Val Phe Arg Leu His Leu Val Leu Gly
            180                 185                 190

atg acc atc ttc ctg ctc ttc ccg ttc acc cga ctg gtg cat gtg tgg     624
Met Thr Ile Phe Leu Leu Phe Pro Phe Thr Arg Leu Val His Val Trp
        195                 200                 205

agc gcg ccg ttt gag tac ttt act cgt cga tat caa att gtg cgt tcg     672
Ser Ala Pro Phe Glu Tyr Phe Thr Arg Arg Tyr Gln Ile Val Arg Ser
    210                 215                 220

cgg cga taa                                                         681
Arg Arg  *
225


<210> 154
<211> 696
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(696)

<400> 154
atg cag atc ctc aaa gtg atc ggc ctg ttg atg gag tat ccg gac gag     48
Met Gln Ile Leu Lys Val Ile Gly Leu Leu Met Glu Tyr Pro Asp Glu
 1               5                  10                  15

ctg ttg tgg gaa tgc aag gag gac gcg ctg gcg ttg atc cgc cgc gac     96
Leu Leu Trp Glu Cys Lys Glu Asp Ala Leu Ala Leu Ile Arg Arg Asp
            20                  25                  30

gcg ccg atg ctt acg gat ttc acc cac aac ctg ctt aac gcg ccg ctg     144
Ala Pro Met Leu Thr Asp Phe Thr His Asn Leu Leu Asn Ala Pro Leu
        35                  40                  45
```

ctg gat aaa cag gcc gaa tgg tgc gaa gtg ttt gac cgc ggg cgc acc 192
Leu Asp Lys Gln Ala Glu Trp Cys Glu Val Phe Asp Arg Gly Arg Thr
50 55 60

acg tcg ctg ctg ctg ttc gaa cat gtt cat gcc gag tcc cgc gat cgc 240
Thr Ser Leu Leu Leu Phe Glu His Val His Ala Glu Ser Arg Asp Arg
65 70 75 80

ggc cag gca atg gtg gac ctg ctg gcg gag tat gaa aag gtc ggc ctg 288
Gly Gln Ala Met Val Asp Leu Leu Ala Glu Tyr Glu Lys Val Gly Leu
85 90 95

cag ctg gat tgt cgg gaa ctg ccc gat tat cta ccg ctg tat ctg gag 336
Gln Leu Asp Cys Arg Glu Leu Pro Asp Tyr Leu Pro Leu Tyr Leu Glu
100 105 110

tat tta agc gta ctg cct gac gat cag gcg aaa gaa gga ttg ctc aat 384
Tyr Leu Ser Val Leu Pro Asp Asp Gln Ala Lys Glu Gly Leu Leu Asn
115 120 125

gtt gcg ccg atc ctc gcc ctg ctt ggc ggt cgc tta aaa caa cgc gag 432
Val Ala Pro Ile Leu Ala Leu Leu Gly Gly Arg Leu Lys Gln Arg Glu
130 135 140

gca ccg tgg tac gcg ttg ttt gat gct ctg ctg caa ctg gcg gga agc 480
Ala Pro Trp Tyr Ala Leu Phe Asp Ala Leu Leu Gln Leu Ala Gly Ser
145 150 155 160

tcc ctt tca agt gac agt gtc aca aaa caa gta aac agc gaa gag cgc 528
Ser Leu Ser Ser Asp Ser Val Thr Lys Gln Val Asn Ser Glu Glu Arg
165 170 175

gat gac acc cgc cag gcg ctt gat gca gtg tgg gaa gag gaa cag gtg 576
Asp Asp Thr Arg Gln Ala Leu Asp Ala Val Trp Glu Glu Glu Gln Val
180 185 190

aag ttt att gaa gat aac gcc acg gcg tgt gac agc tcg ccg ctt aat 624
Lys Phe Ile Glu Asp Asn Ala Thr Ala Cys Asp Ser Ser Pro Leu Asn
195 200 205

caa tat cag cga cgc ttt agc cag gat gtc gcg ccg caa tat gtc gac 672
Gln Tyr Gln Arg Arg Phe Ser Gln Asp Val Ala Pro Gln Tyr Val Asp
210 215 220

atc agt gcg gga ggt ggg aaa tga 696
Ile Ser Ala Gly Gly Gly Lys *
225 230

<210> 155
<211> 1545
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1545)

<400> 155
atg aaa atc cgt tca caa gtc ggc atg gtg ctt aac ctc gac aaa tgt 48

```
Met Lys Ile Arg Ser Gln Val Gly Met Val Leu Asn Leu Asp Lys Cys
 1               5                  10                  15

atc ggc tgc cat acc tgt tcg gtg acc tgt aaa aac gtc tgg acc ggg    96
Ile Gly Cys His Thr Cys Ser Val Thr Cys Lys Asn Val Trp Thr Gly
            20                  25                  30

cgc gaa ggc atg gag tac gca tgg ttt aac aac gtc gaa acc aaa ccg   144
Arg Glu Gly Met Glu Tyr Ala Trp Phe Asn Asn Val Glu Thr Lys Pro
        35                  40                  45

ggc att ggt tat ccg aaa aac tgg gaa gat cag gaa gag tgg caa ggc   192
Gly Ile Gly Tyr Pro Lys Asn Trp Glu Asp Gln Glu Glu Trp Gln Gly
    50                  55                  60

ggc tgg gtg cgt gat gtg aat ggc aag ata cgc ccg cgt ctg ggc aac   240
Gly Trp Val Arg Asp Val Asn Gly Lys Ile Arg Pro Arg Leu Gly Asn
65                  70                  75                  80

aag atg ggc gta ata acc aaa atc ttc gcc aac ccg gtg gtg ccg cag   288
Lys Met Gly Val Ile Thr Lys Ile Phe Ala Asn Pro Val Val Pro Gln
                85                  90                  95

att gat gat tac tac gaa cct ttc acc ttc gac tac gaa cat ttg cat   336
Ile Asp Asp Tyr Tyr Glu Pro Phe Thr Phe Asp Tyr Glu His Leu His
            100                 105                 110

agc gca ccg gaa ggc aaa cat att cct act gct cgc ccg cgt tca ctg   384
Ser Ala Pro Glu Gly Lys His Ile Pro Thr Ala Arg Pro Arg Ser Leu
        115                 120                 125

att gac ggc aaa cgg atg gac aaa gtg atc tgg ggg cca aac tgg gaa   432
Ile Asp Gly Lys Arg Met Asp Lys Val Ile Trp Gly Pro Asn Trp Glu
    130                 135                 140

gaa ctg ctg ggc ggt gag ttc gaa aaa cgt gcc cgc gac cgc aac ttc   480
Glu Leu Leu Gly Gly Glu Phe Glu Lys Arg Ala Arg Asp Arg Asn Phe
145                 150                 155                 160

gag gcc atg caa aag gag atg tac ggg cag ttt gaa aac acc ttc atg   528
Glu Ala Met Gln Lys Glu Met Tyr Gly Gln Phe Glu Asn Thr Phe Met
                165                 170                 175

atg tac ctg ccg cgc ctg tgc gaa cac tgc ctc aat ccc agc tgc gtg   576
Met Tyr Leu Pro Arg Leu Cys Glu His Cys Leu Asn Pro Ser Cys Val
            180                 185                 190

gcg acc tgc cca agc ggc gct atc tac aaa cgc gaa gaa gac ggc att   624
Ala Thr Cys Pro Ser Gly Ala Ile Tyr Lys Arg Glu Glu Asp Gly Ile
        195                 200                 205

gtg ctg att gat cag gat aaa tgc cgt ggc tgg cgt ttg tgc ata agc   672
Val Leu Ile Asp Gln Asp Lys Cys Arg Gly Trp Arg Leu Cys Ile Ser
    210                 215                 220

ggt tgt ccg tac aaa aaa atc tac ttc aac tgg aaa agc ggc aag tca   720
Gly Cys Pro Tyr Lys Lys Ile Tyr Phe Asn Trp Lys Ser Gly Lys Ser
225                 230                 235                 240

gaa aaa tgc atc ttc tgt tac cca cga att gag tcc ggt caa ccg acc   768
Glu Lys Cys Ile Phe Cys Tyr Pro Arg Ile Glu Ser Gly Gln Pro Thr
```

```
                245                 250                 255

gtg tgc tca gaa acc tgc gtg ggt cgc atc cgg tat ctg ggc gtg ctg      816
Val Cys Ser Glu Thr Cys Val Gly Arg Ile Arg Tyr Leu Gly Val Leu
            260                 265                 270

ctt tac gac gcc gac cgc atc gag gaa gcg gcg agc acc gag cgc gaa      864
Leu Tyr Asp Ala Asp Arg Ile Glu Glu Ala Ala Ser Thr Glu Arg Glu
        275                 280                 285

gtt gac ctc tat gaa cgc cag tgc gaa gtg ttc ctc gat cca cac gat      912
Val Asp Leu Tyr Glu Arg Gln Cys Glu Val Phe Leu Asp Pro His Asp
    290                 295                 300

ccc tca gtg atc gag gaa gcc ctg aaa caa ggt att cca caa aat gtg      960
Pro Ser Val Ile Glu Glu Ala Leu Lys Gln Gly Ile Pro Gln Asn Val
305                 310                 315                 320

att gac gct gcc cag cgt tcg cca gtc tac aaa atg gcg atg gac tgg      1008
Ile Asp Ala Ala Gln Arg Ser Pro Val Tyr Lys Met Ala Met Asp Trp
                325                 330                 335

aaa ctg gca cta ccg ttg cac cct gaa tac cgc acc ctg ccg atg gtc      1056
Lys Leu Ala Leu Pro Leu His Pro Glu Tyr Arg Thr Leu Pro Met Val
            340                 345                 350

tgg tac gtt cct ccg ctg tca ccg att cag tcc tac gca gat gcg ggc      1104
Trp Tyr Val Pro Pro Leu Ser Pro Ile Gln Ser Tyr Ala Asp Ala Gly
        355                 360                 365

ggt ttg ccg aaa agc gaa ggc gtg ctg ccc gcc atc gaa agc ctg cgt      1152
Gly Leu Pro Lys Ser Glu Gly Val Leu Pro Ala Ile Glu Ser Leu Arg
    370                 375                 380

att ccg gtg caa tat ctc gcc aat atg ttg agt gcc ggc gat acc ggt      1200
Ile Pro Val Gln Tyr Leu Ala Asn Met Leu Ser Ala Gly Asp Thr Gly
385                 390                 395                 400

ccg gta ctg cgg gcg ctg aaa cgg atg atg gcg atg cgc cac tat atg      1248
Pro Val Leu Arg Ala Leu Lys Arg Met Met Ala Met Arg His Tyr Met
                405                 410                 415

cgt tca caa acc gtg gaa ggc gtt act gat act cgt gcc atc gac gaa      1296
Arg Ser Gln Thr Val Glu Gly Val Thr Asp Thr Arg Ala Ile Asp Glu
            420                 425                 430

gta ggc ctg agc gtc gcc cag gtc gaa gag atg tat cgc tat ctc gcc      1344
Val Gly Leu Ser Val Ala Gln Val Glu Glu Met Tyr Arg Tyr Leu Ala
        435                 440                 445

att gcc aac tat gaa gat cgt ttt gtt atc ccg acc agc cat cgg gaa      1392
Ile Ala Asn Tyr Glu Asp Arg Phe Val Ile Pro Thr Ser His Arg Glu
    450                 455                 460

atg gcg ggc gat gcc ttc gca gaa cgc aac ggc tgc ggt ttt acc ttt      1440
Met Ala Gly Asp Ala Phe Ala Glu Arg Asn Gly Cys Gly Phe Thr Phe
465                 470                 475                 480

ggc gac ggt tgc cac ggt tcg gac agt aaa ttc aac ctg ttc aac agt      1488
Gly Asp Gly Cys His Gly Ser Asp Ser Lys Phe Asn Leu Phe Asn Ser
                485                 490                 495
```

agc cgt atc gat gcc atc aac atc acc gaa gtg cgc gac aaa gcg gag 1536
Ser Arg Ile Asp Ala Ile Asn Ile Thr Glu Val Arg Asp Lys Ala Glu
500 505 510

ggc gaa taa 1545
Gly Glu *

<210> 156
<211> 3741
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(3741)

<400> 156
atg agt aaa ctt ttg gat cgc ttt cgc tac ttc aaa caa aag ggc gaa 48
Met Ser Lys Leu Leu Asp Arg Phe Arg Tyr Phe Lys Gln Lys Gly Glu
1 5 10 15

acc ttt gcc gat ggt cac gga cag gtg atg cat agc aac cgc gac tgg 96
Thr Phe Ala Asp Gly His Gly Gln Val Met His Ser Asn Arg Asp Trp
20 25 30

gag gac agc tat cgc cag cgt tgg cag ttc gac aaa atc gtg cgt tcc 144
Glu Asp Ser Tyr Arg Gln Arg Trp Gln Phe Asp Lys Ile Val Arg Ser
35 40 45

acc cac ggt gtt aac tgt aca ggc tcc tgt agc tgg aaa atc tac gtt 192
Thr His Gly Val Asn Cys Thr Gly Ser Cys Ser Trp Lys Ile Tyr Val
50 55 60

aaa aat ggt ctg gtg acc tgg gaa atc caa cag acc gac tac ccg cgc 240
Lys Asn Gly Leu Val Thr Trp Glu Ile Gln Gln Thr Asp Tyr Pro Arg
65 70 75 80

act cgc cct gac ctg ccc aat cat gaa cct cgc ggc tgc ccg cgt ggc 288
Thr Arg Pro Asp Leu Pro Asn His Glu Pro Arg Gly Cys Pro Arg Gly
85 90 95

gca agt tac tcc tgg tat ctt tac agc gct aac cgc ctg aaa tac ccg 336
Ala Ser Tyr Ser Trp Tyr Leu Tyr Ser Ala Asn Arg Leu Lys Tyr Pro
100 105 110

ctc att cgt aaa cga ctg att gaa ctg tgg cgc gaa gcc ctc aag caa 384
Leu Ile Arg Lys Arg Leu Ile Glu Leu Trp Arg Glu Ala Leu Lys Gln
115 120 125

cac agc gat ccg gta ctg gcg tgg gca tcg att atg aac gat ccg caa 432
His Ser Asp Pro Val Leu Ala Trp Ala Ser Ile Met Asn Asp Pro Gln
130 135 140

aag tgc ctg agc tac aaa caa gtg cgt ggg cgc ggc ggg ttt atc cgc 480
Lys Cys Leu Ser Tyr Lys Gln Val Arg Gly Arg Gly Gly Phe Ile Arg
145 150 155 160

tcc aac tgg cag gaa cta aac cag ctg att gcc gcc gct aac gtc tgg 528

Ser Asn Trp Gln Glu Leu Asn Gln Leu Ile Ala Ala Ala Asn Val Trp
165 170 175

acc atc aaa acc tac ggc ccg gat cgc gtt gcc ggt ttc tcg ccg atc 576
Thr Ile Lys Thr Tyr Gly Pro Asp Arg Val Ala Gly Phe Ser Pro Ile
180 185 190

ccg gcg atg tcg atg gtt tct tac gcc gcc gga acg cgt tat ctg tcg 624
Pro Ala Met Ser Met Val Ser Tyr Ala Ala Gly Thr Arg Tyr Leu Ser
195 200 205

ctg ctt ggc ggc acc tgt tta agt ttc tac gac tgg tat tgc gac ctg 672
Leu Leu Gly Gly Thr Cys Leu Ser Phe Tyr Asp Trp Tyr Cys Asp Leu
210 215 220

ccg ccc gcc tcg ccg atg acc tgg ggc gag caa acc gac gta ccg gaa 720
Pro Pro Ala Ser Pro Met Thr Trp Gly Glu Gln Thr Asp Val Pro Glu
225 230 235 240

tct gcc gac tgg tat aac tcc agc tac atc atc gcc tgg ggg tct aac 768
Ser Ala Asp Trp Tyr Asn Ser Ser Tyr Ile Ile Ala Trp Gly Ser Asn
245 250 255

gta ccg cag aca cgt acg ccg gac gcc cac ttc ttt acc gaa gta cgc 816
Val Pro Gln Thr Arg Thr Pro Asp Ala His Phe Phe Thr Glu Val Arg
260 265 270

tac aaa ggc act aaa acc atc gcc att acc cct gac tac tct gaa gtg 864
Tyr Lys Gly Thr Lys Thr Ile Ala Ile Thr Pro Asp Tyr Ser Glu Val
275 280 285

gcc aaa ttg tgc gac cag tgg ctg gca ccg aaa caa ggc act gat agc 912
Ala Lys Leu Cys Asp Gln Trp Leu Ala Pro Lys Gln Gly Thr Asp Ser
290 295 300

gcc ctg gcg atg gca atg ggc cat gtg att tta aaa gag ttt cat ctc 960
Ala Leu Ala Met Ala Met Gly His Val Ile Leu Lys Glu Phe His Leu
305 310 315 320

gat aat ccc agc gac tac ttt atc aac tac tgc cgc cgc tac agc gac 1008
Asp Asn Pro Ser Asp Tyr Phe Ile Asn Tyr Cys Arg Arg Tyr Ser Asp
325 330 335

atg ccg atg ctg gta atg ctg gag cct cgc gac gat ggt agc tac gtt 1056
Met Pro Met Leu Val Met Leu Glu Pro Arg Asp Asp Gly Ser Tyr Val
340 345 350

ccc ggg cgc atg atc cgc gca tct gac ctg gtg gat gga ctg ggc gaa 1104
Pro Gly Arg Met Ile Arg Ala Ser Asp Leu Val Asp Gly Leu Gly Glu
355 360 365

agc aac aat ccg cag tgg aaa acc gta gca gtt aat acc gca ggt gaa 1152
Ser Asn Asn Pro Gln Trp Lys Thr Val Ala Val Asn Thr Ala Gly Glu
370 375 380

ttg gta gtg ccg aac ggt tcg att ggt ttc cgc tgg gga gaa aaa ggc 1200
Leu Val Val Pro Asn Gly Ser Ile Gly Phe Arg Trp Gly Glu Lys Gly
385 390 395 400

aaa tgg aat ctg gaa tcc att gcc gcc ggt acg gaa acc gaa ttg tcg 1248
Lys Trp Asn Leu Glu Ser Ile Ala Ala Gly Thr Glu Thr Glu Leu Ser

```
                405                 410                 415

tta acc ctg ctc ggt caa cat gac gct gtt gca ggc gtg gcc ttc ccc    1296
Leu Thr Leu Leu Gly Gln His Asp Ala Val Ala Gly Val Ala Phe Pro
            420                 425                 430

tac ttt ggc ggc att gaa aat ccg cat ttt cgc agc gta aaa cac aat    1344
Tyr Phe Gly Gly Ile Glu Asn Pro His Phe Arg Ser Val Lys His Asn
        435                 440                 445

ccg gtg ctg gtg cgc caa ttg ccc gtt aaa aac ctg acg tta gtc gat    1392
Pro Val Leu Val Arg Gln Leu Pro Val Lys Asn Leu Thr Leu Val Asp
    450                 455                 460

ggc aac acc tgt ccg gtg gtc agc gtt tat gat ttg gta ctg gcg aat    1440
Gly Asn Thr Cys Pro Val Val Ser Val Tyr Asp Leu Val Leu Ala Asn
465                 470                 475                 480

tac ggc ctc gat cgc ggg ctg gaa gat gaa aac agt gcg aaa gat tac    1488
Tyr Gly Leu Asp Arg Gly Leu Glu Asp Glu Asn Ser Ala Lys Asp Tyr
                485                 490                 495

gct gaa atc aaa ccg tac acc cca gcc tgg ggt gag caa att acc ggc    1536
Ala Glu Ile Lys Pro Tyr Thr Pro Ala Trp Gly Glu Gln Ile Thr Gly
            500                 505                 510

gtg ccg cgc cag tat att gaa acc atc gcc cgt gaa ttt gcc gat acc    1584
Val Pro Arg Gln Tyr Ile Glu Thr Ile Ala Arg Glu Phe Ala Asp Thr
        515                 520                 525

gcc cat aaa acg cat ggg cgc tcg atg att atc ctc ggc gca ggt gtt    1632
Ala His Lys Thr His Gly Arg Ser Met Ile Ile Leu Gly Ala Gly Val
    530                 535                 540

aac cac tgg tat cac atg gac atg aac tac cgc ggg atg atc aat atg    1680
Asn His Trp Tyr His Met Asp Met Asn Tyr Arg Gly Met Ile Asn Met
545                 550                 555                 560

ctg atc ttc tgc ggt tgt gtc ggg caa agc ggt ggc ggc tgg gca cac    1728
Leu Ile Phe Cys Gly Cys Val Gly Gln Ser Gly Gly Gly Trp Ala His
                565                 570                 575

tat gtc ggc cag gaa aaa ctg cgc cca caa acc ggc tgg ttg cca ctg    1776
Tyr Val Gly Gln Glu Lys Leu Arg Pro Gln Thr Gly Trp Leu Pro Leu
            580                 585                 590

gcc ttt gcg ctc gac tgg aac cga cca ccg cgc caa atg aac agc acc    1824
Ala Phe Ala Leu Asp Trp Asn Arg Pro Pro Arg Gln Met Asn Ser Thr
        595                 600                 605

tcg ttt ttc tac aat cat tcc agc caa tgg cgc tat gaa aaa gtc tct    1872
Ser Phe Phe Tyr Asn His Ser Ser Gln Trp Arg Tyr Glu Lys Val Ser
    610                 615                 620

gct cag gag tta ctt tca ccg ctc gcc gat gcc agt aag tac agc ggt    1920
Ala Gln Glu Leu Leu Ser Pro Leu Ala Asp Ala Ser Lys Tyr Ser Gly
625                 630                 635                 640

cat ctg att gat ttc aac gtt cgc gcc gaa cgt atg ggc tgg cta cct    1968
His Leu Ile Asp Phe Asn Val Arg Ala Glu Arg Met Gly Trp Leu Pro
                645                 650                 655
```

```
tct gcg ccg cag ctg ggg cgt aac ccg ctc ggg att aaa gct gaa gcc    2016
Ser Ala Pro Gln Leu Gly Arg Asn Pro Leu Gly Ile Lys Ala Glu Ala
            660                 665                 670

gac aag gcc gga tta tcc ccc aca gaa ttt acc gcc cag gcg ctg aaa    2064
Asp Lys Ala Gly Leu Ser Pro Thr Glu Phe Thr Ala Gln Ala Leu Lys
        675                 680                 685

tcg ggc gat tta cgt atg gcc tgc gaa caa cca gat agc agc agc aat    2112
Ser Gly Asp Leu Arg Met Ala Cys Glu Gln Pro Asp Ser Ser Ser Asn
    690                 695                 700

cat ccg cgt aat ttg ttt gtc tgg cgc tct aac ctg ctt ggc tcc tcc    2160
His Pro Arg Asn Leu Phe Val Trp Arg Ser Asn Leu Leu Gly Ser Ser
705                 710                 715                 720

ggc aaa ggc cac gag tat atg cag aag tat ctg ctg ggg acc gaa agc    2208
Gly Lys Gly His Glu Tyr Met Gln Lys Tyr Leu Leu Gly Thr Glu Ser
                725                 730                 735

ggg att cag ggc gag gaa ctc ggt gcc agc gac ggg atc aaa ccg gaa    2256
Gly Ile Gln Gly Glu Glu Leu Gly Ala Ser Asp Gly Ile Lys Pro Glu
            740                 745                 750

gaa gtc gag tgg caa act gca gcg att gaa ggc aag ctc gac ctg ctg    2304
Glu Val Glu Trp Gln Thr Ala Ala Ile Glu Gly Lys Leu Asp Leu Leu
        755                 760                 765

gtg acg ctc gac ttc cgc atg tcc agt acc tgc ctg ttc tcc gat atc    2352
Val Thr Leu Asp Phe Arg Met Ser Ser Thr Cys Leu Phe Ser Asp Ile
    770                 775                 780

gtt ctg ccc acc gcc acc tgg tac gaa aaa gac gat atg aac acc tcg    2400
Val Leu Pro Thr Ala Thr Trp Tyr Glu Lys Asp Asp Met Asn Thr Ser
785                 790                 795                 800

gat atg cat ccg ttt att cat ccg ctt tct gcg gcg gtc gat ccg gcg    2448
Asp Met His Pro Phe Ile His Pro Leu Ser Ala Ala Val Asp Pro Ala
                805                 810                 815

tgg gaa tca cgc agc gac tgg gaa atc tac aaa ggt atc gcc aaa gca    2496
Trp Glu Ser Arg Ser Asp Trp Glu Ile Tyr Lys Gly Ile Ala Lys Ala
            820                 825                 830

ttt tcg caa gtg tgc gtg ggt cat ctt ggc aaa gaa acc gac gtg gta    2544
Phe Ser Gln Val Cys Val Gly His Leu Gly Lys Glu Thr Asp Val Val
        835                 840                 845

tta caa ccc cta ctg cat gac tct ccg gca gag ctc tca cag ccg tgt    2592
Leu Gln Pro Leu Leu His Asp Ser Pro Ala Glu Leu Ser Gln Pro Cys
    850                 855                 860

gaa gtg ctc gac tgg cgc aaa ggc gaa tgc gat ctg att cca ggc aaa    2640
Glu Val Leu Asp Trp Arg Lys Gly Glu Cys Asp Leu Ile Pro Gly Lys
865                 870                 875                 880

acc gcg ccg aat att gtg gcg gtg gag cgc gac tac cct gct acg tat    2688
Thr Ala Pro Asn Ile Val Ala Val Glu Arg Asp Tyr Pro Ala Thr Tyr
                885                 890                 895
```

gaa cgc ttt acc tcg ctc ggg cca ttg atg gac aaa ctt ggc aat ggc 2736
Glu Arg Phe Thr Ser Leu Gly Pro Leu Met Asp Lys Leu Gly Asn Gly
900 905 910

ggt aaa ggg att tcg tgg aat acc cag gat gaa atc gat ttc ctc ggc 2784
Gly Lys Gly Ile Ser Trp Asn Thr Gln Asp Glu Ile Asp Phe Leu Gly
915 920 925

aaa ctc aat tac acc aag cgt gat ggc ccg gca cag ggg cgt ccg ctg 2832
Lys Leu Asn Tyr Thr Lys Arg Asp Gly Pro Ala Gln Gly Arg Pro Leu
930 935 940

att gac acc gcc att gac gct tca gaa gtg att ctg gca ctg gca ccg 2880
Ile Asp Thr Ala Ile Asp Ala Ser Glu Val Ile Leu Ala Leu Ala Pro
945 950 955 960

gaa acc aac ggt cat gtt gca gtc aaa gcg tgg cag gcg ctg ggc gag 2928
Glu Thr Asn Gly His Val Ala Val Lys Ala Trp Gln Ala Leu Gly Glu
965 970 975

atc acc ggg cgc gaa cat acc cat ctg gcg ctg cac aaa gag gac gag 2976
Ile Thr Gly Arg Glu His Thr His Leu Ala Leu His Lys Glu Asp Glu
980 985 990

aag att cgc ttt cgc gat att cag gcg cag ccg cgt aaa att atc tcc 3024
Lys Ile Arg Phe Arg Asp Ile Gln Ala Gln Pro Arg Lys Ile Ile Ser
995 1000 1005

agc ccc acc tgg tcc ggt ctg gaa agc gat cac gtc tcc tat aac gcg 3072
Ser Pro Thr Trp Ser Gly Leu Glu Ser Asp His Val Ser Tyr Asn Ala
1010 1015 1020

gga tac acc aac gtt cat gag tta att ccg tgg cgc acg ctg tcg gga 3120
Gly Tyr Thr Asn Val His Glu Leu Ile Pro Trp Arg Thr Leu Ser Gly
1025 1030 1035 1040

cgc cag cag ctc tat cag gat cat ccg tgg atg cgt gct ttt ggt gaa 3168
Arg Gln Gln Leu Tyr Gln Asp His Pro Trp Met Arg Ala Phe Gly Glu
1045 1050 1055

agc ctg gtg gct tat cgc ccg cct atc gac acc cgt agc gtc agt gag 3216
Ser Leu Val Ala Tyr Arg Pro Pro Ile Asp Thr Arg Ser Val Ser Glu
1060 1065 1070

atg cgc cag ata ccg cca aac ggc ttc ccg gaa aaa gca ctt aac ttc 3264
Met Arg Gln Ile Pro Pro Asn Gly Phe Pro Glu Lys Ala Leu Asn Phe
1075 1080 1085

ctg acg ccg cac cag aaa tgg ggc att cac tca acc tac agt gaa aac 3312
Leu Thr Pro His Gln Lys Trp Gly Ile His Ser Thr Tyr Ser Glu Asn
1090 1095 1100

ctg cta atg ctg acg ctc tct cgc ggt gga ccg att gtc tgg atc agc 3360
Leu Leu Met Leu Thr Leu Ser Arg Gly Gly Pro Ile Val Trp Ile Ser
1105 1110 1115 1120

gaa aca gat gcc cga gaa ctg acc att gtc gat aac gac tgg gtg gaa 3408
Glu Thr Asp Ala Arg Glu Leu Thr Ile Val Asp Asn Asp Trp Val Glu
1125 1130 1135

gta ttc aac gcc aac ggc gcg ctg act gcc cgc gcg gtg gtc agc caa 3456

Val Phe Asn Ala Asn Gly Ala Leu Thr Ala Arg Ala Val Val Ser Gln
1140 1145 1150

cgt gta ccg ccg ggc atg acc atg atg tat cac gcc cag gaa cgc att 3504
Arg Val Pro Pro Gly Met Thr Met Met Tyr His Ala Gln Glu Arg Ile
1155 1160 1165

atg aat att cct ggt tcg gaa gta act ggc atg cgc ggc ggc att cat 3552
Met Asn Ile Pro Gly Ser Glu Val Thr Gly Met Arg Gly Gly Ile His
1170 1175 1180

aac tcg gtt acc cgc gtt tgc ccg aaa cca acg cat atg att ggc ggt 3600
Asn Ser Val Thr Arg Val Cys Pro Lys Pro Thr His Met Ile Gly Gly
1185 1190 1195 1200

tac gcg cag ctg gcc tgg ggc ttt aac tac tac ggc acc gtc gga tcg 3648
Tyr Ala Gln Leu Ala Trp Gly Phe Asn Tyr Tyr Gly Thr Val Gly Ser
1205 1210 1215

aac cgc gat gag ttc atc atg atc cgc aag atg aag aac gtt aac tgg 3696
Asn Arg Asp Glu Phe Ile Met Ile Arg Lys Met Lys Asn Val Asn Trp
1220 1225 1230

ctg gat gat gaa ggt cgc gat cag gta cag gag gcg aaa aaa tga 3741
Leu Asp Asp Glu Gly Arg Asp Gln Val Gln Glu Ala Lys Lys *
1235 1240 1245

<210> 157
<211> 498
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(498)

<400> 157
atg gat ttg tca cag cta aca cca cgt cgt ccc tat ctg ctg cgt gca 48
Met Asp Leu Ser Gln Leu Thr Pro Arg Arg Pro Tyr Leu Leu Arg Ala
1 5 10 15

ttc tat gag tgg ttg ctg gat aac cag ctc acg ccg cac ctg gtg gtg 96
Phe Tyr Glu Trp Leu Leu Asp Asn Gln Leu Thr Pro His Leu Val Val
20 25 30

gat gtg acg ctc cct ggc gtg cag gtt cct atg gaa tat gcg cgt gac 144
Asp Val Thr Leu Pro Gly Val Gln Val Pro Met Glu Tyr Ala Arg Asp
35 40 45

ggg caa atc gta ctc aac att gcg ccg cgt gct gtc ggc aat ctg gaa 192
Gly Gln Ile Val Leu Asn Ile Ala Pro Arg Ala Val Gly Asn Leu Glu
50 55 60

ctg gcg aat gat gag gtg cgc ttt aac gcg cgc ttt ggt ggc att ccg 240
Leu Ala Asn Asp Glu Val Arg Phe Asn Ala Arg Phe Gly Gly Ile Pro
65 70 75 80

cgt cag gtt tct gtg ccg ctg gct gcc gtg ctg gct atc tac gcc cgt 288
Arg Gln Val Ser Val Pro Leu Ala Ala Val Leu Ala Ile Tyr Ala Arg

```
                  85                  90                  95

gaa aat ggc gca ggc acg atg ttt gag cct gaa gct gcc tac gat gaa    336
Glu Asn Gly Ala Gly Thr Met Phe Glu Pro Glu Ala Ala Tyr Asp Glu
            100                 105                 110

gat acc agc atc atg aat gat gaa gag gca tcg gca gac aac gaa acc    384
Asp Thr Ser Ile Met Asn Asp Glu Glu Ala Ser Ala Asp Asn Glu Thr
        115                 120                 125

gtt atg tcg gtt att gat ggc gac aag cca gat cac gat gat gac act    432
Val Met Ser Val Ile Asp Gly Asp Lys Pro Asp His Asp Asp Asp Thr
    130                 135                 140

cat cct gac gat gaa cct ccg cag cca cca cgc ggt ggt cga ccg gca    480
His Pro Asp Asp Glu Pro Pro Gln Pro Pro Arg Gly Gly Arg Pro Ala
145                 150                 155                 160

tta cgc gtt gtg aag taa                                            498
Leu Arg Val Val Lys  *
                165


&lt;210&gt; 158
&lt;211&gt; 639
&lt;212&gt; DNA
&lt;213&gt; Escherichia coli

&lt;220&gt;
&lt;221&gt; CDS
&lt;222&gt; (1)...(639)

&lt;400&gt; 158
atg gct gtc gct gcc aac aaa cgt tcg gta atg acg ctg ttt tcc ggt     48
Met Ala Val Ala Ala Asn Lys Arg Ser Val Met Thr Leu Phe Ser Gly
 1               5                  10                  15

cct act gac atc tat agc cat cag gtc cgc att gtg ctg gct gag aaa     96
Pro Thr Asp Ile Tyr Ser His Gln Val Arg Ile Val Leu Ala Glu Lys
            20                  25                  30

ggt gta agt ttc gag atc gaa cac gtg gaa aag gac aat ccg cct cag    144
Gly Val Ser Phe Glu Ile Glu His Val Glu Lys Asp Asn Pro Pro Gln
        35                  40                  45

gat ctg att gac ctc aac ccg aat cag agc gtt ccg acc ctg gtg gat    192
Asp Leu Ile Asp Leu Asn Pro Asn Gln Ser Val Pro Thr Leu Val Asp
    50                  55                  60

cgt gag ctg acc ctg tgg gaa tct cgc atc att atg gaa tat ctg gat    240
Arg Glu Leu Thr Leu Trp Glu Ser Arg Ile Ile Met Glu Tyr Leu Asp
65                  70                  75                  80

gag cgt ttc ccg cat ccg cca ctg atg cct gtt tac ccg gta gct cgc    288
Glu Arg Phe Pro His Pro Pro Leu Met Pro Val Tyr Pro Val Ala Arg
                85                  90                  95

ggt gaa agc cgt ctg tac atg cat cgc atc gaa aaa gac tgg tac acg    336
Gly Glu Ser Arg Leu Tyr Met His Arg Ile Glu Lys Asp Trp Tyr Thr
            100                 105                 110
```

ctg atg aac acc atc atc aac ggt tca gct tct gaa gca gat gcc gca 384
Leu Met Asn Thr Ile Ile Asn Gly Ser Ala Ser Glu Ala Asp Ala Ala
115 120 125

cgt aag caa ctg cgc gaa gaa ctg ctg gcg att gcg ccg gtc ttc ggt 432
Arg Lys Gln Leu Arg Glu Glu Leu Leu Ala Ile Ala Pro Val Phe Gly
130 135 140

cag aag ccg tac ttc ctg agc gat gag ttc agc ctg gtc gat tgc tat 480
Gln Lys Pro Tyr Phe Leu Ser Asp Glu Phe Ser Leu Val Asp Cys Tyr
145 150 155 160

ctt gct ccg ctg ctg tgg cgt ctg ccg caa ctg ggc atc gag ttc agc 528
Leu Ala Pro Leu Leu Trp Arg Leu Pro Gln Leu Gly Ile Glu Phe Ser
165 170 175

ggc ccg ggt gcg aaa gag ctg aaa ggc tat atg acc cgc gtc ttt gag 576
Gly Pro Gly Ala Lys Glu Leu Lys Gly Tyr Met Thr Arg Val Phe Glu
180 185 190

cgt gac tct ttc ctt gct tct tta act gaa gca gaa cgt gaa atg cgt 624
Arg Asp Ser Phe Leu Ala Ser Leu Thr Glu Ala Glu Arg Glu Met Arg
195 200 205

ctg ggc cgg agt taa 639
Leu Gly Arg Ser *
210

<210> 159
<211> 1518
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1518)

<400> 159
atg tct gaa caa cac gca cag ggc gct gac gcg gta gtc gat ctt aac 48
Met Ser Glu Gln His Ala Gln Gly Ala Asp Ala Val Val Asp Leu Asn
1 5 10 15

aat gaa ctg aaa acg cgt cgt gag aag ctg gcg aac ctg cgc gag cag 96
Asn Glu Leu Lys Thr Arg Arg Glu Lys Leu Ala Asn Leu Arg Glu Gln
20 25 30

ggg att gcc ttc ccg aac gat ttc cgt cgc gat cat acc tct gac caa 144
Gly Ile Ala Phe Pro Asn Asp Phe Arg Arg Asp His Thr Ser Asp Gln
35 40 45

ttg cac gca gaa ttc gac ggc aaa gag aac gaa gaa ctg gaa gcg ctg 192
Leu His Ala Glu Phe Asp Gly Lys Glu Asn Glu Glu Leu Glu Ala Leu
50 55 60

aac atc gaa gtc gcc gtt gct ggc cgc atg atg acc cgt cgt att atg 240
Asn Ile Glu Val Ala Val Ala Gly Arg Met Met Thr Arg Arg Ile Met
65 70 75 80

ggt aaa gcg tct ttc gtt acc ctg cag gac gtt ggc ggt cgc att cag 288
Gly Lys Ala Ser Phe Val Thr Leu Gln Asp Val Gly Gly Arg Ile Gln

85 90 95

ctg tac gtt gcc cgt gac gat ctc ccg gaa ggc gtt tat aac gag cag 336
Leu Tyr Val Ala Arg Asp Asp Leu Pro Glu Gly Val Tyr Asn Glu Gln
100 105 110

ttc aaa aaa tgg gac ctc ggc gac atc ctc ggc gcg aaa ggt aag ctg 384
Phe Lys Lys Trp Asp Leu Gly Asp Ile Leu Gly Ala Lys Gly Lys Leu
115 120 125

ttc aaa acc aaa acc ggc gaa ctg tct atc cac tgc acc gag ttg cgt 432
Phe Lys Thr Lys Thr Gly Glu Leu Ser Ile His Cys Thr Glu Leu Arg
130 135 140

ctg ctg acc aaa gca ctg cgt ccg ctg ccg gat aaa ttc cac ggc ttg 480
Leu Leu Thr Lys Ala Leu Arg Pro Leu Pro Asp Lys Phe His Gly Leu
145 150 155 160

cag gat cag gaa gcg cgc tat cgt cag cgt tat ctc gat ctc atc tcc 528
Gln Asp Gln Glu Ala Arg Tyr Arg Gln Arg Tyr Leu Asp Leu Ile Ser
165 170 175

aac gat gaa tcc cgc aac acc ttt aaa gtg cgc tcg cag atc ctc tct 576
Asn Asp Glu Ser Arg Asn Thr Phe Lys Val Arg Ser Gln Ile Leu Ser
180 185 190

ggt att cgc cag ttc atg gtg aac cgc ggc ttt atg gaa gtt gaa acg 624
Gly Ile Arg Gln Phe Met Val Asn Arg Gly Phe Met Glu Val Glu Thr
195 200 205

ccg atg atg cag gtg atc cct ggc ggt gcc gct gcg cgt ccg ttt atc 672
Pro Met Met Gln Val Ile Pro Gly Gly Ala Ala Ala Arg Pro Phe Ile
210 215 220

acc cac cat aac gcg ctg gat ctc gac atg tac ctg cgt atc gcg ccg 720
Thr His His Asn Ala Leu Asp Leu Asp Met Tyr Leu Arg Ile Ala Pro
225 230 235 240

gaa ctg tac ctc aag cgt ctg gtg gtt ggt ggc ttc gag cgt gta ttc 768
Glu Leu Tyr Leu Lys Arg Leu Val Val Gly Gly Phe Glu Arg Val Phe
245 250 255

gaa atc aac cgt aac ttc cgt aac gaa ggt att tcc gta cgt cat aac 816
Glu Ile Asn Arg Asn Phe Arg Asn Glu Gly Ile Ser Val Arg His Asn
260 265 270

cca gag ttc acc atg atg gaa ctc tac atg gct tac gca gat tac aaa 864
Pro Glu Phe Thr Met Met Glu Leu Tyr Met Ala Tyr Ala Asp Tyr Lys
275 280 285

gat ctg atc gag ctg acc gaa tcg ctg ttc cgt act ctg gca cag gat 912
Asp Leu Ile Glu Leu Thr Glu Ser Leu Phe Arg Thr Leu Ala Gln Asp
290 295 300

att ctc ggt aag acg gaa gtg acc tac ggc gac gtg acg ctg gac ttc 960
Ile Leu Gly Lys Thr Glu Val Thr Tyr Gly Asp Val Thr Leu Asp Phe
305 310 315 320

ggt aaa ccg ttc gaa aaa ctg acc atg cgt gaa gcg atc aag aaa tat 1008
Gly Lys Pro Phe Glu Lys Leu Thr Met Arg Glu Ala Ile Lys Lys Tyr
325 330 335

cgc ccg gaa acc gac atg gcg gat ctg gac aac ttc gac tct gcg aaa 1056
Arg Pro Glu Thr Asp Met Ala Asp Leu Asp Asn Phe Asp Ser Ala Lys
340 345 350

gca att gct gaa tct atc ggc atc cac gtt gag aag agc tgg ggt ctg 1104
Ala Ile Ala Glu Ser Ile Gly Ile His Val Glu Lys Ser Trp Gly Leu
355 360 365

ggc cgt atc gtt acc gag atc ttc gaa gaa gtg gca gaa gca cat ctg 1152
Gly Arg Ile Val Thr Glu Ile Phe Glu Glu Val Ala Glu Ala His Leu
370 375 380

att cag ccg acc ttc att act gaa tat ccg gca gaa gtt tct ccg ctg 1200
Ile Gln Pro Thr Phe Ile Thr Glu Tyr Pro Ala Glu Val Ser Pro Leu
385 390 395 400

gcg cgt cgt aac gac gtt aac ccg gaa atc aca gac cgc ttt gag ttc 1248
Ala Arg Arg Asn Asp Val Asn Pro Glu Ile Thr Asp Arg Phe Glu Phe
405 410 415

ttc att ggt ggt cgt gaa atc ggt aac ggc ttt agc gag ctg aat gac 1296
Phe Ile Gly Gly Arg Glu Ile Gly Asn Gly Phe Ser Glu Leu Asn Asp
420 425 430

gcg gaa gat cag gcg caa cgc ttc ctg gat cag gtt gcc gcg aaa gac 1344
Ala Glu Asp Gln Ala Gln Arg Phe Leu Asp Gln Val Ala Ala Lys Asp
435 440 445

gca ggt gac gac gaa gcg atg ttc tac gat gaa gat tac gtc acc gca 1392
Ala Gly Asp Asp Glu Ala Met Phe Tyr Asp Glu Asp Tyr Val Thr Ala
450 455 460

ctg gaa cat ggc tta ccg ccg aca gca ggt ctg gga att ggt atc gac 1440
Leu Glu His Gly Leu Pro Pro Thr Ala Gly Leu Gly Ile Gly Ile Asp
465 470 475 480

cgt atg gta atg ctg ttc acc aac agc cat acc atc cgc gac gtt att 1488
Arg Met Val Met Leu Phe Thr Asn Ser His Thr Ile Arg Asp Val Ile
485 490 495

ctg ttc ccg gcg atg cgt ccg gta aaa taa 1518
Leu Phe Pro Ala Met Arg Pro Val Lys *
500 505

<210> 160
<211> 1098
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1098)

<400> 160
atg ttt gaa att aat ccg gta aat aat cgc att cag gac ctc acg gaa 48
Met Phe Glu Ile Asn Pro Val Asn Asn Arg Ile Gln Asp Leu Thr Glu
1 5 10 15

cgc tcc gac gtt ctt agg ggg tat ctt gac tac gac gcc aag aaa gag 96

```
Arg Ser Asp Val Leu Arg Gly Tyr Leu Asp Tyr Asp Ala Lys Lys Glu
            20                  25                  30

cgt ctg gaa gaa gta aac gcc gag ctg gaa cag ccg gat gtc tgg aac      144
Arg Leu Glu Glu Val Asn Ala Glu Leu Glu Gln Pro Asp Val Trp Asn
        35                  40                  45

gaa ccc gaa cgc gca cag gcg ctg ggt aaa gag cgt tcc tcc ctc gaa      192
Glu Pro Glu Arg Ala Gln Ala Leu Gly Lys Glu Arg Ser Ser Leu Glu
    50                  55                  60

gcc gtt gtc gac acc ctc gac caa atg aaa cag ggg ctg gaa gat gtt      240
Ala Val Val Asp Thr Leu Asp Gln Met Lys Gln Gly Leu Glu Asp Val
65                  70                  75                  80

tct ggt ctg ctg gaa ctg gct gta gaa gct gac gac gaa gaa acc ttt      288
Ser Gly Leu Leu Glu Leu Ala Val Glu Ala Asp Asp Glu Glu Thr Phe
                85                  90                  95

aac gaa gcc gtt gct gaa ctc gac gcc ctg gaa gaa aaa ctg gcg cag      336
Asn Glu Ala Val Ala Glu Leu Asp Ala Leu Glu Glu Lys Leu Ala Gln
            100                 105                 110

ctt gag ttc cgc cgt atg ttc tct ggc gaa tat gac agc gcc gac tgc      384
Leu Glu Phe Arg Arg Met Phe Ser Gly Glu Tyr Asp Ser Ala Asp Cys
        115                 120                 125

tac ctc gat att cag gcg ggg tct ggc ggt acg gaa gca cag gac tgg      432
Tyr Leu Asp Ile Gln Ala Gly Ser Gly Gly Thr Glu Ala Gln Asp Trp
    130                 135                 140

gcg agc atg ctt gag cgt atg tat ctg cgc tgg gca gaa tcg cgt ggt      480
Ala Ser Met Leu Glu Arg Met Tyr Leu Arg Trp Ala Glu Ser Arg Gly
145                 150                 155                 160

ttc aaa act gaa atc atc gaa gag tcg gaa ggt gaa gtg gcg ggt att      528
Phe Lys Thr Glu Ile Ile Glu Glu Ser Glu Gly Glu Val Ala Gly Ile
                165                 170                 175

aaa tcc gtg acg atc aaa atc tcc ggc gat tac gct tac ggc tgg ctg      576
Lys Ser Val Thr Ile Lys Ile Ser Gly Asp Tyr Ala Tyr Gly Trp Leu
            180                 185                 190

cgt aca gaa acc ggc gtt cac cgc ctg gtg cgt aaa agc ccg ttt gac      624
Arg Thr Glu Thr Gly Val His Arg Leu Val Arg Lys Ser Pro Phe Asp
        195                 200                 205

tcc ggc ggt cgt cgc cac acg tcg ttc agc tcc gcg ttt gtt tat ccg      672
Ser Gly Gly Arg Arg His Thr Ser Phe Ser Ser Ala Phe Val Tyr Pro
    210                 215                 220

gaa gtt gat gat gat att gat atc gaa atc aac ccg gcg gat ctg cgc      720
Glu Val Asp Asp Asp Ile Asp Ile Glu Ile Asn Pro Ala Asp Leu Arg
225                 230                 235                 240

att gac gtt tat cgc acg tcc ggc gcg ggc ggt cag cac gtt aac cgt      768
Ile Asp Val Tyr Arg Thr Ser Gly Ala Gly Gly Gln His Val Asn Arg
                245                 250                 255

acc gaa tct gcg gtg cgt att acc cac atc ccg acc ggg atc gtg acc      816
Thr Glu Ser Ala Val Arg Ile Thr His Ile Pro Thr Gly Ile Val Thr
```

```
            260                 265                 270

cag tgc cag aac gac cgt tcc cag cac aag aac aaa gat cag gcc atg    864
Gln Cys Gln Asn Asp Arg Ser Gln His Lys Asn Lys Asp Gln Ala Met
        275                 280                 285

aag cag atg aaa gcg aag ctt tat gaa ctg gag atg cag aag aaa aat    912
Lys Gln Met Lys Ala Lys Leu Tyr Glu Leu Glu Met Gln Lys Lys Asn
    290                 295                 300

gcc gag aaa cag gcg atg gaa gat aac aaa tcc gac atc ggc tgg ggc    960
Ala Glu Lys Gln Ala Met Glu Asp Asn Lys Ser Asp Ile Gly Trp Gly
305                 310                 315                 320

agc cag att cgt tct tat gtc ctt gat gac tcc cgc att aaa gat ctg    1008
Ser Gln Ile Arg Ser Tyr Val Leu Asp Asp Ser Arg Ile Lys Asp Leu
                325                 330                 335

cgc acc ggg gta gaa acc cgc aac acg cag gcc gtg ctg gac ggc agc    1056
Arg Thr Gly Val Glu Thr Arg Asn Thr Gln Ala Val Leu Asp Gly Ser
            340                 345                 350

ctg gat caa ttt atc gaa gca agt ttg aaa gca ggg tta tga            1098
Leu Asp Gln Phe Ile Glu Ala Ser Leu Lys Ala Gly Leu  *
        355                 360                 365
```

<210> 161
<211> 1734
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1734)

<400> 161

```
gtg aaa caa cag ata caa ctt cgt cgc cgt gaa gtc gat gaa acg gca    48
Met Lys Gln Gln Ile Gln Leu Arg Arg Arg Glu Val Asp Glu Thr Ala
 1               5                  10                  15

gac ttg ccc gct gaa ttg cct ccc ttg ctg cgc cgt tta tac gcc agc    96
Asp Leu Pro Ala Glu Leu Pro Pro Leu Leu Arg Arg Leu Tyr Ala Ser
            20                  25                  30

cgg gga gta cgc agt gcg caa gaa ctg gaa cgc agt gtt aaa ggt atg    144
Arg Gly Val Arg Ser Ala Gln Glu Leu Glu Arg Ser Val Lys Gly Met
        35                  40                  45

ctg ccc tgg cag caa ctg agc ggc gtc gaa aag gcc gtt gag atc ctt    192
Leu Pro Trp Gln Gln Leu Ser Gly Val Glu Lys Ala Val Glu Ile Leu
    50                  55                  60

tac aac gct ttt cgc gaa gga acg cgg att att gtg gtc ggt gat ttc    240
Tyr Asn Ala Phe Arg Glu Gly Thr Arg Ile Ile Val Val Gly Asp Phe
65                  70                  75                  80

gac gcc gac ggc gcg acc agc acg gct cta agc gtg ctg gcg atg cgc    288
Asp Ala Asp Gly Ala Thr Ser Thr Ala Leu Ser Val Leu Ala Met Arg
                85                  90                  95
```

```
tcg ctt ggt tgc agc aat atc gac tac ctg gta cca aac cgt ttc gaa    336
Ser Leu Gly Cys Ser Asn Ile Asp Tyr Leu Val Pro Asn Arg Phe Glu
            100                 105                 110

gac ggt tac ggc tta agc ccg gaa gtg gtc gat cag gcc cat gcc cgt    384
Asp Gly Tyr Gly Leu Ser Pro Glu Val Val Asp Gln Ala His Ala Arg
        115                 120                 125

ggc gcg cag tta att gtc acg gtg gat aac ggt att tcc tcc cat gcg    432
Gly Ala Gln Leu Ile Val Thr Val Asp Asn Gly Ile Ser Ser His Ala
    130                 135                 140

ggg gtt gag cac gct cgc tcg ttg ggc atc ccg gtt att gtt acc gat    480
Gly Val Glu His Ala Arg Ser Leu Gly Ile Pro Val Ile Val Thr Asp
145                 150                 155                 160

cac cat ttg cca ggc gac aca tta ccc gca gcg gaa gcg atc att aac    528
His His Leu Pro Gly Asp Thr Leu Pro Ala Ala Glu Ala Ile Ile Asn
                165                 170                 175

cct aac ttg cgc gac tgt aat ttc ccg tcg aaa tca ctg gca ggc gtg    576
Pro Asn Leu Arg Asp Cys Asn Phe Pro Ser Lys Ser Leu Ala Gly Val
            180                 185                 190

ggt gtg gcg ttt tat ctg atg ctg gcg ctg cgc acc ttt ttg cgc gat    624
Gly Val Ala Phe Tyr Leu Met Leu Ala Leu Arg Thr Phe Leu Arg Asp
        195                 200                 205

cag ggc tgg ttt gat gag cgt aac atc gca att cct aac ctg gca gaa    672
Gln Gly Trp Phe Asp Glu Arg Asn Ile Ala Ile Pro Asn Leu Ala Glu
    210                 215                 220

ctg ctg gat ctg gtc gcg ctg ggg aca gtg gcg gac gtc gtg ccg ctg    720
Leu Leu Asp Leu Val Ala Leu Gly Thr Val Ala Asp Val Val Pro Leu
225                 230                 235                 240

gac gct aat aat cgc att ctg acc tgg cag ggg atg agt cgc atc cga    768
Asp Ala Asn Asn Arg Ile Leu Thr Trp Gln Gly Met Ser Arg Ile Arg
                245                 250                 255

gcc gga aag tgc cgt ccg ggg att aaa gcg ctg ctt gaa gtg gca aac    816
Ala Gly Lys Cys Arg Pro Gly Ile Lys Ala Leu Leu Glu Val Ala Asn
            260                 265                 270

cgt gat gca caa aaa ctc gcc gcc agc gat tta ggt ttt gcg ctg ggg    864
Arg Asp Ala Gln Lys Leu Ala Ala Ser Asp Leu Gly Phe Ala Leu Gly
        275                 280                 285

cca cgt ctc aat gct gcc gga cga ctg gac gat atg tcc gtc ggt gtg    912
Pro Arg Leu Asn Ala Ala Gly Arg Leu Asp Asp Met Ser Val Gly Val
    290                 295                 300

gcg ctg ttg ttg tgc gac aac atc ggc gaa gcg cgc gtg ctg gca aat    960
Ala Leu Leu Leu Cys Asp Asn Ile Gly Glu Ala Arg Val Leu Ala Asn
305                 310                 315                 320

gaa ctc gat gcg cta aac cag acg cga aaa gag atc gaa caa gga atg   1008
Glu Leu Asp Ala Leu Asn Gln Thr Arg Lys Glu Ile Glu Gln Gly Met
                325                 330                 335
```

```
caa att gaa gcc ctg acc ctg tgc gag aaa ctg gag cgc agc cgt gac     1056
Gln Ile Glu Ala Leu Thr Leu Cys Glu Lys Leu Glu Arg Ser Arg Asp
            340                 345                 350

acg cta ccc ggc ggg ctg gca atg tat cac ccc gaa tgg cat cag ggc     1104
Thr Leu Pro Gly Gly Leu Ala Met Tyr His Pro Glu Trp His Gln Gly
        355                 360                 365

gtt gtc ggt att ctg gct tcg cgc atc aaa gag cgt ttt cac cgt ccg     1152
Val Val Gly Ile Leu Ala Ser Arg Ile Lys Glu Arg Phe His Arg Pro
    370                 375                 380

gtt atc gcg ttt gcg cca gca ggt gac ggt acg ctg aaa ggt tcc ggt     1200
Val Ile Ala Phe Ala Pro Ala Gly Asp Gly Thr Leu Lys Gly Ser Gly
385                 390                 395                 400

cgc tcc att cag ggg ctg cat atg cgt gat gcg ctg gag cga tta gac     1248
Arg Ser Ile Gln Gly Leu His Met Arg Asp Ala Leu Glu Arg Leu Asp
                405                 410                 415

aca ctc tac cct ggc atg atg ctg aag ttt ggc ggt cat gcg atg gcg     1296
Thr Leu Tyr Pro Gly Met Met Leu Lys Phe Gly Gly His Ala Met Ala
            420                 425                 430

gcg ggt ttg tcg ctg gaa gag gat aaa ttc aaa ctc ttt caa caa cgg     1344
Ala Gly Leu Ser Leu Glu Glu Asp Lys Phe Lys Leu Phe Gln Gln Arg
        435                 440                 445

ttt ggc gaa ctg gtt act gag tgg ctg gac cct tcg cta ttg caa ggc     1392
Phe Gly Glu Leu Val Thr Glu Trp Leu Asp Pro Ser Leu Leu Gln Gly
    450                 455                 460

gaa gtg gta tca gac ggt ccg tta agc ccg gcc gaa atg acc atg gaa     1440
Glu Val Val Ser Asp Gly Pro Leu Ser Pro Ala Glu Met Thr Met Glu
465                 470                 475                 480

gtg gcg cag ctg ctg cgc gat gct ggc ccg tgg ggg cag atg ttc ccg     1488
Val Ala Gln Leu Leu Arg Asp Ala Gly Pro Trp Gly Gln Met Phe Pro
                485                 490                 495

gag ccg ctg ttt gac ggt cat ttc cgt ctg ctg caa cag cgg ctg gtg     1536
Glu Pro Leu Phe Asp Gly His Phe Arg Leu Leu Gln Gln Arg Leu Val
            500                 505                 510

ggc gaa cgt cat ttg aag gtg atg gtc gaa ccg gtc ggc ggc ggt cca     1584
Gly Glu Arg His Leu Lys Val Met Val Glu Pro Val Gly Gly Gly Pro
        515                 520                 525

ctg ctg gat ggt att gct ttt aat gtc gat acc gcc ctc tgg ccg gat     1632
Leu Leu Asp Gly Ile Ala Phe Asn Val Asp Thr Ala Leu Trp Pro Asp
    530                 535                 540

aac ggc gtg cgc gaa gtg caa ctg gct tat aag ctc gat atc aac gag     1680
Asn Gly Val Arg Glu Val Gln Leu Ala Tyr Lys Leu Asp Ile Asn Glu
545                 550                 555                 560

ttt cgc ggc aac cgc agc ctg caa att atc atc gac aat atc tgg cca     1728
Phe Arg Gly Asn Arg Ser Leu Gln Ile Ile Ile Asp Asn Ile Trp Pro
                565                 570                 575

att tag                                                             1734
```

Ile *

<210> 162
<211> 711
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(711)

<400> 162
atg aag aaa ggt ttt atg ttg ttt act ttg tta gcg gcg ttt tca ggc 48
Met Lys Lys Gly Phe Met Leu Phe Thr Leu Leu Ala Ala Phe Ser Gly
1 5 10 15

ttt gct cag gct gat gac gcg gca att caa caa acg tta gcc aaa atg 96
Phe Ala Gln Ala Asp Asp Ala Ala Ile Gln Gln Thr Leu Ala Lys Met
20 25 30

ggc atc aaa agc agc gat att cag ccc gcg cct gta gct ggc atg aag 144
Gly Ile Lys Ser Ser Asp Ile Gln Pro Ala Pro Val Ala Gly Met Lys
35 40 45

aca gtt ctg act aac agc ggc gtg ttg tac atc acc gat gat ggt aaa 192
Thr Val Leu Thr Asn Ser Gly Val Leu Tyr Ile Thr Asp Asp Gly Lys
50 55 60

cat atc att cag ggg cca atg tat gac gtt agt ggc acg gct ccg gtc 240
His Ile Ile Gln Gly Pro Met Tyr Asp Val Ser Gly Thr Ala Pro Val
65 70 75 80

aat gtc acc aat aag atg ctg tta aag cag ttg aat gcg ctt gaa aaa 288
Asn Val Thr Asn Lys Met Leu Leu Lys Gln Leu Asn Ala Leu Glu Lys
85 90 95

gag atg atc gtt tat aaa gcg ccg cag gaa aaa cac gtc atc acc gtg 336
Glu Met Ile Val Tyr Lys Ala Pro Gln Glu Lys His Val Ile Thr Val
100 105 110

ttt act gat att acc tgt ggt tac tgc cac aaa ctg cat gag caa atg 384
Phe Thr Asp Ile Thr Cys Gly Tyr Cys His Lys Leu His Glu Gln Met
115 120 125

gca gac tac aac gcg ctg ggg atc acc gtg cgt tat ctt gct ttc ccg 432
Ala Asp Tyr Asn Ala Leu Gly Ile Thr Val Arg Tyr Leu Ala Phe Pro
130 135 140

cgc cag ggg ctg gac agc gat gca gag aaa gaa atg aaa gct atc tgg 480
Arg Gln Gly Leu Asp Ser Asp Ala Glu Lys Glu Met Lys Ala Ile Trp
145 150 155 160

tgt gcg aaa gat aaa aac aaa gcg ttt gat gat gtg atg gca ggt aaa 528
Cys Ala Lys Asp Lys Asn Lys Ala Phe Asp Asp Val Met Ala Gly Lys
165 170 175

agc gtc gca cca gcc agt tgc gac gtg gat att gcc gac cat tac gca 576
Ser Val Ala Pro Ala Ser Cys Asp Val Asp Ile Ala Asp His Tyr Ala
180 185 190

ctt ggc gtc cag ctt ggc gtt agc ggt act ccg gca gtt gtg ctg agc 624
Leu Gly Val Gln Leu Gly Val Ser Gly Thr Pro Ala Val Val Leu Ser
195 200 205

aat ggc aca ctt gtt ccg ggt tac cag ccg ccg aaa gag atg aaa gaa 672
Asn Gly Thr Leu Val Pro Gly Tyr Gln Pro Pro Lys Glu Met Lys Glu
210 215 220

ttc ctc gac gaa cac caa aaa atg acc agc ggt aaa taa 711
Phe Leu Asp Glu His Gln Lys Met Thr Ser Gly Lys *
225 230 235

<210> 163
<211> 897
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(897)

<400> 163
gtg aaa cag gat ctg gca cgc atc gag cag ttt ctt gat gct ctg tgg 48
Met Lys Gln Asp Leu Ala Arg Ile Glu Gln Phe Leu Asp Ala Leu Trp
1 5 10 15

ctg gaa aaa aat ctg gct gaa aat acg ttg aac gct tac cgt cgc gat 96
Leu Glu Lys Asn Leu Ala Glu Asn Thr Leu Asn Ala Tyr Arg Arg Asp
20 25 30

ctg tca atg atg gtg gag tgg ttg cat cac cgc ggg ttg acg ctg gcg 144
Leu Ser Met Met Val Glu Trp Leu His His Arg Gly Leu Thr Leu Ala
35 40 45

acg gcg caa agt gac gat ttg cag gca tta ctg gca gaa cgg ctg gag 192
Thr Ala Gln Ser Asp Asp Leu Gln Ala Leu Leu Ala Glu Arg Leu Glu
50 55 60

ggc ggg tat aaa gcc acc agc tca gcg cgt ttg ctg agt gca gtg cgc 240
Gly Gly Tyr Lys Ala Thr Ser Ser Ala Arg Leu Leu Ser Ala Val Arg
65 70 75 80

cga ttg ttc cag tat ctt tat cgc gaa aag ttt cgt gaa gac gat ccc 288
Arg Leu Phe Gln Tyr Leu Tyr Arg Glu Lys Phe Arg Glu Asp Asp Pro
85 90 95

agt gcg cat ctc gct tca ccg aaa ttg ccc cag cgt ttg cca aaa gat 336
Ser Ala His Leu Ala Ser Pro Lys Leu Pro Gln Arg Leu Pro Lys Asp
100 105 110

tta agc gaa gcg cag gtc gaa cgt tta tta cag gca cca tta atc gat 384
Leu Ser Glu Ala Gln Val Glu Arg Leu Leu Gln Ala Pro Leu Ile Asp
115 120 125

cag cca ctg gag cta cgc gat aaa gcc atg ctt gaa gtg ttg tat gct 432
Gln Pro Leu Glu Leu Arg Asp Lys Ala Met Leu Glu Val Leu Tyr Ala
130 135 140

acc ggg ctg cgt gtc tct gaa ctg gtc gga ctg aca atg agt gat atc 480

Thr Gly Leu Arg Val Ser Glu Leu Val Gly Leu Thr Met Ser Asp Ile
145 150 155 160

agc ctg cgt cag ggc gtg gta cgg gtc att ggt aaa ggc aac aaa gag 528
Ser Leu Arg Gln Gly Val Val Arg Val Ile Gly Lys Gly Asn Lys Glu
165 170 175

cgt ctg gtg ccg tta ggt gaa gag gcg gtt tac tgg ctg gaa acc tat 576
Arg Leu Val Pro Leu Gly Glu Glu Ala Val Tyr Trp Leu Glu Thr Tyr
180 185 190

ctg gaa cat ggg cgt ccg tgg ctg ctg aat ggt gtg tca att gac gtg 624
Leu Glu His Gly Arg Pro Trp Leu Leu Asn Gly Val Ser Ile Asp Val
195 200 205

ttg ttt ccc agc cag cgt gcg cag cag atg acg cga cag acc ttc tgg 672
Leu Phe Pro Ser Gln Arg Ala Gln Gln Met Thr Arg Gln Thr Phe Trp
210 215 220

cac cgt att aaa cat tat gct gtg ctg gcg ggt atc gac agc gaa aag 720
His Arg Ile Lys His Tyr Ala Val Leu Ala Gly Ile Asp Ser Glu Lys
225 230 235 240

ctg tca ccg cat gtg ttg cgt cac gct ttt gcc act cat tta tta aat 768
Leu Ser Pro His Val Leu Arg His Ala Phe Ala Thr His Leu Leu Asn
245 250 255

cat ggt gcg gat tta cgc gtg gtg cag atg cta ctg ggc cac agc gat 816
His Gly Ala Asp Leu Arg Val Val Gln Met Leu Leu Gly His Ser Asp
260 265 270

ctc tcc acc acg caa att tat acg cat gtc gct acc gag cgt ctg cgg 864
Leu Ser Thr Thr Gln Ile Tyr Thr His Val Ala Thr Glu Arg Leu Arg
275 280 285

caa ctt cat caa cag cat cac ccg cgg gcg tga 897
Gln Leu His Gln Gln His His Pro Arg Ala *
290 295

<210> 164
<211> 1104
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1104)

<400> 164
atg agt gac agc cag acg ctg gtg gta aaa ctc ggc acc agt gtg cta 48
Met Ser Asp Ser Gln Thr Leu Val Val Lys Leu Gly Thr Ser Val Leu
1 5 10 15

aca ggc gga tcg cgc cgt ctg aac cgt gcc cat atc gtt gaa ctt gtt 96
Thr Gly Gly Ser Arg Arg Leu Asn Arg Ala His Ile Val Glu Leu Val
20 25 30

cgc cag tgc gcg cag tta cat gcc gcc ggg cat cgg att gtt att gtg 144
Arg Gln Cys Ala Gln Leu His Ala Ala Gly His Arg Ile Val Ile Val
35 40 45

```
acg tcg ggc gcg atc gcc gcc gga cgt gag cac ctg ggt tac ccg gaa    192
Thr Ser Gly Ala Ile Ala Ala Gly Arg Glu His Leu Gly Tyr Pro Glu
     50                  55                  60

ctg cca gcg acc atc gcc tcg aaa caa ctg ctg gcg gcg gta ggg cag    240
Leu Pro Ala Thr Ile Ala Ser Lys Gln Leu Leu Ala Ala Val Gly Gln
 65                  70                  75                  80

agt cga ctg att caa ctg tgg gaa cag ctg ttt tcg att tat ggc att    288
Ser Arg Leu Ile Gln Leu Trp Glu Gln Leu Phe Ser Ile Tyr Gly Ile
                 85                  90                  95

cac gtc ggg caa atg ctg ctg acc cgt gct gat atg gaa gac cgt gaa    336
His Val Gly Gln Met Leu Leu Thr Arg Ala Asp Met Glu Asp Arg Glu
            100                 105                 110

cgc ttc ctg aac gcc cgc gac acc ctg cga gcg ttg ctc gat aac aat    384
Arg Phe Leu Asn Ala Arg Asp Thr Leu Arg Ala Leu Leu Asp Asn Asn
        115                 120                 125

atc gtt ccg gta atc aat gag aac gat gct gtc gct acg gca gag att    432
Ile Val Pro Val Ile Asn Glu Asn Asp Ala Val Ala Thr Ala Glu Ile
    130                 135                 140

aag gtc ggc gat aac gat aac ctt tct gcg ctg gcg gcg att ctt gcg    480
Lys Val Gly Asp Asn Asp Asn Leu Ser Ala Leu Ala Ala Ile Leu Ala
145                 150                 155                 160

ggt gcc gat aaa ctg ttg ctg ctg acc gat caa aaa ggt ttg tat acc    528
Gly Ala Asp Lys Leu Leu Leu Leu Thr Asp Gln Lys Gly Leu Tyr Thr
                165                 170                 175

gct gac ccg cgc agc aat ccg cag gca gaa ctg att aaa gat gtt tac    576
Ala Asp Pro Arg Ser Asn Pro Gln Ala Glu Leu Ile Lys Asp Val Tyr
            180                 185                 190

ggc att gat gac gca ctg cgc gcg att gcc ggt gac agc gtt tca ggc    624
Gly Ile Asp Asp Ala Leu Arg Ala Ile Ala Gly Asp Ser Val Ser Gly
        195                 200                 205

ctc gga act ggc ggc atg agt acc aaa ttg cag gcc gct gac gtg gct    672
Leu Gly Thr Gly Gly Met Ser Thr Lys Leu Gln Ala Ala Asp Val Ala
    210                 215                 220

tgc cgt gcg ggt atc gac acc att att gcc gcg ggc agc aag ccg ggc    720
Cys Arg Ala Gly Ile Asp Thr Ile Ile Ala Ala Gly Ser Lys Pro Gly
225                 230                 235                 240

gtt att ggt gat gtg atg gaa ggc att tcc gtc ggt acg ctg ttc cat    768
Val Ile Gly Asp Val Met Glu Gly Ile Ser Val Gly Thr Leu Phe His
                245                 250                 255

gcc cag gcg act ccg ctt gaa aac cgt aaa cgc tgg att ttc ggt gcg    816
Ala Gln Ala Thr Pro Leu Glu Asn Arg Lys Arg Trp Ile Phe Gly Ala
            260                 265                 270

ccg ccg gcg ggt gaa atc acg gta gat gaa ggg gca act gcc gcc att    864
Pro Pro Ala Gly Glu Ile Thr Val Asp Glu Gly Ala Thr Ala Ala Ile
        275                 280                 285
```

```
ctg gaa cgc ggc agc tcc ctg ttg ccg aaa ggc att aaa agc gtg act    912
Leu Glu Arg Gly Ser Ser Leu Leu Pro Lys Gly Ile Lys Ser Val Thr
    290                 295                 300

ggc aat ttc tcg cgt ggt gaa gtc atc cgc att tgc aac ctc gaa ggc    960
Gly Asn Phe Ser Arg Gly Glu Val Ile Arg Ile Cys Asn Leu Glu Gly
305                 310                 315                 320

cgc gat atc gcc cac ggc gtc agt cgt tac aac agc gat gca tta cgc    1008
Arg Asp Ile Ala His Gly Val Ser Arg Tyr Asn Ser Asp Ala Leu Arg
                325                 330                 335

cgt att gcc gga cac cac tcg caa gaa att gat gca ata ctg gga tat    1056
Arg Ile Ala Gly His His Ser Gln Glu Ile Asp Ala Ile Leu Gly Tyr
            340                 345                 350

gaa tac ggc ccg gtt gcc gtt cac cgt gat gac atg att acc cgt taa    1104
Glu Tyr Gly Pro Val Ala Val His Arg Asp Asp Met Ile Thr Arg  *
        355                 360                 365
```

<210> 165
<211> 1254
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1254)

<400> 165

```
atg ctg gaa caa atg ggc att gcc gcg aag caa gcc tcg tat aaa tta    48
Met Leu Glu Gln Met Gly Ile Ala Ala Lys Gln Ala Ser Tyr Lys Leu
 1               5                  10                  15

gcg caa ctc tcc agc cgc gaa aaa aat cgc gtg ctg gaa aaa atc gcc    96
Ala Gln Leu Ser Ser Arg Glu Lys Asn Arg Val Leu Glu Lys Ile Ala
            20                  25                  30

gat gaa ctg gaa gca caa agc gaa atc atc ctc aac gct aac gcc cag    144
Asp Glu Leu Glu Ala Gln Ser Glu Ile Ile Leu Asn Ala Asn Ala Gln
        35                  40                  45

gat gtt gct gac gcg cga gcc aat ggc ctt agc gaa gcg atg ctt gac    192
Asp Val Ala Asp Ala Arg Ala Asn Gly Leu Ser Glu Ala Met Leu Asp
    50                  55                  60

cgt ctg gca ctg acg ccc gca cgg ctg aaa ggc att gcc gac gat gta    240
Arg Leu Ala Leu Thr Pro Ala Arg Leu Lys Gly Ile Ala Asp Asp Val
65                  70                  75                  80

cgt cag gtg tgc aac ctc gcc gat ccg gtg ggg cag gta atc gat ggc    288
Arg Gln Val Cys Asn Leu Ala Asp Pro Val Gly Gln Val Ile Asp Gly
                85                  90                  95

ggc gta ctg gac agc ggc ctg cgt ctt gag cgt cgt cgc gta ccg ctg    336
Gly Val Leu Asp Ser Gly Leu Arg Leu Glu Arg Arg Arg Val Pro Leu
            100                 105                 110

ggg gtt att ggc gtg att tat gaa gcg cgc ccg aac gtg acg gtt gat    384
```

Gly Val Ile Gly Val Ile Tyr Glu Ala Arg Pro Asn Val Thr Val Asp
115 120 125

gtc gct tcg ctg tgc ctg aaa acc ggt aat gcg gtg atc ctg cgc ggt 432
Val Ala Ser Leu Cys Leu Lys Thr Gly Asn Ala Val Ile Leu Arg Gly
130 135 140

ggc aaa gaa acg tgt cgc act aac gct gca acg gtg gcg gtg att cag 480
Gly Lys Glu Thr Cys Arg Thr Asn Ala Ala Thr Val Ala Val Ile Gln
145 150 155 160

gac gcc ctg aaa tcc tgc ggc tta ccg gcg ggt gcc gtg cag gcg att 528
Asp Ala Leu Lys Ser Cys Gly Leu Pro Ala Gly Ala Val Gln Ala Ile
165 170 175

gat aat cct gac cgt gcg ctg gtc agt gaa atg ctg cgt atg gat aaa 576
Asp Asn Pro Asp Arg Ala Leu Val Ser Glu Met Leu Arg Met Asp Lys
180 185 190

tac atc gac atg ctg atc ccg cgt ggt ggc gct ggt ttg cat aaa ctg 624
Tyr Ile Asp Met Leu Ile Pro Arg Gly Gly Ala Gly Leu His Lys Leu
195 200 205

tgc cgt gaa cag tcg aca atc ccg gtg atc aca ggt ggt ata ggc gta 672
Cys Arg Glu Gln Ser Thr Ile Pro Val Ile Thr Gly Gly Ile Gly Val
210 215 220

tgc cat att tac gtt gat gaa agt gta gag atc gct gaa gca tta aaa 720
Cys His Ile Tyr Val Asp Glu Ser Val Glu Ile Ala Glu Ala Leu Lys
225 230 235 240

gtg atc gtc aac gcg aaa act cag cgt ccg agc aca tgt aat acg gtt 768
Val Ile Val Asn Ala Lys Thr Gln Arg Pro Ser Thr Cys Asn Thr Val
245 250 255

gaa acg ttg ctg gtg aat aaa aac atc gcc gat agc ttc ctg ccc gca 816
Glu Thr Leu Leu Val Asn Lys Asn Ile Ala Asp Ser Phe Leu Pro Ala
260 265 270

tta agc aaa caa atg gcg gaa agc ggc gtg aca tta cac gca gat gca 864
Leu Ser Lys Gln Met Ala Glu Ser Gly Val Thr Leu His Ala Asp Ala
275 280 285

gct gca ctg gcg cag ttg cag gca ggc cct gcg aag gtg gtt gct gtt 912
Ala Ala Leu Ala Gln Leu Gln Ala Gly Pro Ala Lys Val Val Ala Val
290 295 300

aaa gcc gaa gag tat gac gat gag ttt ctg tca tta gat ttg aac gtc 960
Lys Ala Glu Glu Tyr Asp Asp Glu Phe Leu Ser Leu Asp Leu Asn Val
305 310 315 320

aaa atc gtc agc gat ctt gac gat gcc atc gcc cat att cgt gaa cac 1008
Lys Ile Val Ser Asp Leu Asp Asp Ala Ile Ala His Ile Arg Glu His
325 330 335

ggc aca caa cac tcc gat gcg atc ctg acc cgc gat atg cgc aac gcc 1056
Gly Thr Gln His Ser Asp Ala Ile Leu Thr Arg Asp Met Arg Asn Ala
340 345 350

cag cgt ttt gtt aac gaa gtg gat tcg tcc gct gtt tac gtt aac gcc 1104
Gln Arg Phe Val Asn Glu Val Asp Ser Ser Ala Val Tyr Val Asn Ala

355 360 365

tct acg cgt ttt acc gac ggc ggc cag ttt ggt ctg ggt gcg gaa gtg 1152
Ser Thr Arg Phe Thr Asp Gly Gly Gln Phe Gly Leu Gly Ala Glu Val
370 375 380

gcg gta agc aca caa aaa ctc cac gcg cgt ggc cca atg ggg ctg gaa 1200
Ala Val Ser Thr Gln Lys Leu His Ala Arg Gly Pro Met Gly Leu Glu
385 390 395 400

gca ctg acc act tac aag tgg atc ggc att ggt gat tac acc att cgt 1248
Ala Leu Thr Thr Tyr Lys Trp Ile Gly Ile Gly Asp Tyr Thr Ile Arg
405 410 415

gcg taa 1254
Ala *

<210> 166
<211> 561
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(561)

<400> 166
atg atg aca agg cag gca tca atg aaa ggc ttc cca att gcg cat att 48
Met Met Thr Arg Gln Ala Ser Met Lys Gly Phe Pro Ile Ala His Ile
1 5 10 15

ttt cac cct tca atc ccg cca atg cac gca gtg gtt aac aat cac aat 96
Phe His Pro Ser Ile Pro Pro Met His Ala Val Val Asn Asn His Asn
20 25 30

aga aat att gat tat tgg acg gta aaa aga aag ttt gca gaa att gtc 144
Arg Asn Ile Asp Tyr Trp Thr Val Lys Arg Lys Phe Ala Glu Ile Val
35 40 45

tcc acc aat gac gtt aat aaa att tac agt ata agt aat gaa ctg cgg 192
Ser Thr Asn Asp Val Asn Lys Ile Tyr Ser Ile Ser Asn Glu Leu Arg
50 55 60

aga gta tta tct gca ata act gca ttg aat ttc tat cat ggc gat gtt 240
Arg Val Leu Ser Ala Ile Thr Ala Leu Asn Phe Tyr His Gly Asp Val
65 70 75 80

cct tct gtc atg atc cga atc caa ccg gaa aat atg agt cca ttc att 288
Pro Ser Val Met Ile Arg Ile Gln Pro Glu Asn Met Ser Pro Phe Ile
85 90 95

ata gat att tct aca gga gaa cat gat gat tat atc ata caa aca tta 336
Ile Asp Ile Ser Thr Gly Glu His Asp Asp Tyr Ile Ile Gln Thr Leu
100 105 110

gat gta ggc act ttt gca cct ttt ggt gaa caa tgt act tgc tca gcc 384
Asp Val Gly Thr Phe Ala Pro Phe Gly Glu Gln Cys Thr Cys Ser Ala
115 120 125

```
gtc aat aaa aaa gag ctg gaa tgt att aag gag acg att tct aag tat     432
Val Asn Lys Lys Glu Leu Glu Cys Ile Lys Glu Thr Ile Ser Lys Tyr
    130                 135                 140

tgt gca aaa ttc acc cga aaa gaa gcc atc tta acc ccc ctt gtg cac     480
Cys Ala Lys Phe Thr Arg Lys Glu Ala Ile Leu Thr Pro Leu Val His
145                 150                 155                 160

ttt aat aaa acc agc att act tca gat tgt tgg caa att ctc ttt ttt     528
Phe Asn Lys Thr Ser Ile Thr Ser Asp Cys Trp Gln Ile Leu Phe Phe
                165                 170                 175

tca ccc gat cat ttt aat aat gat ttt tat tga                         561
Ser Pro Asp His Phe Asn Asn Asp Phe Tyr  *
            180                 185


<210> 167
<211> 1188
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1188)

<400> 167
atg ttc cct gtt tca tca att ggt aat gat att agt agt gat tta gtc     48
Met Phe Pro Val Ser Ser Ile Gly Asn Asp Ile Ser Ser Asp Leu Val
 1               5                  10                  15

cgt cga aaa atg aat gac ctt cct gaa agc cca aca gga aat aac ctc     96
Arg Arg Lys Met Asn Asp Leu Pro Glu Ser Pro Thr Gly Asn Asn Leu
            20                  25                  30

gaa gca cta gca cca ggt ata gaa aaa cta aaa cag acc tct att gaa     144
Glu Ala Leu Ala Pro Gly Ile Glu Lys Leu Lys Gln Thr Ser Ile Glu
        35                  40                  45

atg gtc act tta ctt aac acg tta caa cct ggt gga aaa tgc att atc     192
Met Val Thr Leu Leu Asn Thr Leu Gln Pro Gly Gly Lys Cys Ile Ile
    50                  55                  60

act ggt gat ttt caa aaa gaa tta gcg tac tta caa aat gta att ctt     240
Thr Gly Asp Phe Gln Lys Glu Leu Ala Tyr Leu Gln Asn Val Ile Leu
65                  70                  75                  80

tat aat gtc tcg tct ctt cgt ctg gat ttt tta ggt tat aac gcc caa     288
Tyr Asn Val Ser Ser Leu Arg Leu Asp Phe Leu Gly Tyr Asn Ala Gln
                85                  90                  95

att att caa cga tcg gac aat act tgt gaa ctt acc att aat gaa ccg     336
Ile Ile Gln Arg Ser Asp Asn Thr Cys Glu Leu Thr Ile Asn Glu Pro
            100                 105                 110

tta aaa aac cag gaa ata tcc aca ggt aat atc aat att aat tgc cca     384
Leu Lys Asn Gln Glu Ile Ser Thr Gly Asn Ile Asn Ile Asn Cys Pro
        115                 120                 125

tta aaa gat att tac aat gaa atc agg agg tta aac gta att ttt agt     432
Leu Lys Asp Ile Tyr Asn Glu Ile Arg Arg Leu Asn Val Ile Phe Ser
```

130 135 140

tgt ggg act gga gat atc gtt gat cta tcc tct ctg gac tta cgt aat 480
Cys Gly Thr Gly Asp Ile Val Asp Leu Ser Ser Leu Asp Leu Arg Asn
145 150 155 160

gtc gat tta gat tat tat gat ttc aca gat aaa cat atg gct aat act 528
Val Asp Leu Asp Tyr Tyr Asp Phe Thr Asp Lys His Met Ala Asn Thr
165 170 175

att tta aat cct ttt aaa ttg aat tca aca aat ttt act aat gcc aac 576
Ile Leu Asn Pro Phe Lys Leu Asn Ser Thr Asn Phe Thr Asn Ala Asn
180 185 190

atg ttt cag gtt aat ttt gtt agt tca aca caa aac gcc aca atc tcc 624
Met Phe Gln Val Asn Phe Val Ser Ser Thr Gln Asn Ala Thr Ile Ser
195 200 205

tgg gat tat tta cta aaa ata acg cct gtt tta ata agc att agc gat 672
Trp Asp Tyr Leu Leu Lys Ile Thr Pro Val Leu Ile Ser Ile Ser Asp
210 215 220

atg tat tct gaa gaa aaa atc aag ttt gtc gaa agt tgt tta aat gag 720
Met Tyr Ser Glu Glu Lys Ile Lys Phe Val Glu Ser Cys Leu Asn Glu
225 230 235 240

cct gga gac att acc gaa gaa caa tta aaa att atg aga ttt gca att 768
Pro Gly Asp Ile Thr Glu Glu Gln Leu Lys Ile Met Arg Phe Ala Ile
245 250 255

ata aaa tct ata cca agg gca act ctt aca gat aaa tta gaa aat gaa 816
Ile Lys Ser Ile Pro Arg Ala Thr Leu Thr Asp Lys Leu Glu Asn Glu
260 265 270

tta aca aaa gaa ata tat aaa agc tca tcg aaa atc atc aat tgc ttg 864
Leu Thr Lys Glu Ile Tyr Lys Ser Ser Ser Lys Ile Ile Asn Cys Leu
275 280 285

aac aga att aaa tta aca gag atg aaa gaa ttc tca tca gaa aaa ata 912
Asn Arg Ile Lys Leu Thr Glu Met Lys Glu Phe Ser Ser Glu Lys Ile
290 295 300

tat gat tac atc gat ata atc att gaa gat tat gaa aat acc aaa gaa 960
Tyr Asp Tyr Ile Asp Ile Ile Ile Glu Asp Tyr Glu Asn Thr Lys Glu
305 310 315 320

aat gct tat ctg gtc gtc ccc caa att aat tat act atg gat tta aac 1008
Asn Ala Tyr Leu Val Val Pro Gln Ile Asn Tyr Thr Met Asp Leu Asn
325 330 335

ata gaa gac tct agc tca gaa gag tta ctt tca gat aat acc ctc gag 1056
Ile Glu Asp Ser Ser Ser Glu Glu Leu Leu Ser Asp Asn Thr Leu Glu
340 345 350

aaa gac gaa aat tct ccg gac aat ggc ttt gag gtc ggg gaa tat aac 1104
Lys Asp Glu Asn Ser Pro Asp Asn Gly Phe Glu Val Gly Glu Tyr Asn
355 360 365

aca tat gaa gca tat aac tca gag aag caa tat ttt acc aga gag gac 1152
Thr Tyr Glu Ala Tyr Asn Ser Glu Lys Gln Tyr Phe Thr Arg Glu Asp
370 375 380

tat acg tat gat tac gac ctt tta aat gca ata tag 1188
Tyr Thr Tyr Asp Tyr Asp Leu Leu Asn Ala Ile *
385 390 395

<210> 168
<211> 948
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(948)

<400> 168
gtg tgt cat cgc gca ttt cga ctt cat ctt tgc aag gac tgg gtt ttc 48
Met Cys His Arg Ala Phe Arg Leu His Leu Cys Lys Asp Trp Val Phe
1 5 10 15

atg ttt tct ggg ctg tta atc att ctg gtt ccc ctg att gtg ggt tac 96
Met Phe Ser Gly Leu Leu Ile Ile Leu Val Pro Leu Ile Val Gly Tyr
20 25 30

ctc att ccg ctt cgc caa caa gct gcg tta aaa gtt att aat cag cta 144
Leu Ile Pro Leu Arg Gln Gln Ala Ala Leu Lys Val Ile Asn Gln Leu
35 40 45

tta agc tgg atg gtt tac ctt att ctc ttt ttt atg ggt atc agt ctg 192
Leu Ser Trp Met Val Tyr Leu Ile Leu Phe Phe Met Gly Ile Ser Leu
50 55 60

gcg ttt ctc gat aac ctc gcc agt aac ctg ttg gcg att ctg cat tat 240
Ala Phe Leu Asp Asn Leu Ala Ser Asn Leu Leu Ala Ile Leu His Tyr
65 70 75 80

tct gcc gtc agt att acc gtt att tta ctg tgt aat att gcc gcc ctg 288
Ser Ala Val Ser Ile Thr Val Ile Leu Leu Cys Asn Ile Ala Ala Leu
85 90 95

atg tgg ctg gag cga ggc ctg ccg tgg cgc aac cac cat cag caa gaa 336
Met Trp Leu Glu Arg Gly Leu Pro Trp Arg Asn His His Gln Gln Glu
100 105 110

aaa ctc ccg tcg cgt att gcg atg gcg ctg gag tcg cta aaa ctg tgc 384
Lys Leu Pro Ser Arg Ile Ala Met Ala Leu Glu Ser Leu Lys Leu Cys
115 120 125

ggc gta gta gtg att ggt ttt gcc att ggt cta agt gga ctg gct ttc 432
Gly Val Val Val Ile Gly Phe Ala Ile Gly Leu Ser Gly Leu Ala Phe
130 135 140

tta caa cac gcg acc gaa gcc agt gaa tac acg tta att ttg cta ctt 480
Leu Gln His Ala Thr Glu Ala Ser Glu Tyr Thr Leu Ile Leu Leu Leu
145 150 155 160

ttc ctc gtt ggt att cag ttg cgc aat aat ggc atg acc tta aag cag 528
Phe Leu Val Gly Ile Gln Leu Arg Asn Asn Gly Met Thr Leu Lys Gln
165 170 175

att gtc ctt aat cgc cgg gga atg att gtc gcc gtg gtg gtg gtt gtc 576

```
Ile Val Leu Asn Arg Arg Gly Met Ile Val Ala Val Val Val Val Val
        180                 185                 190

agt tca tta att ggt ggt tta att aac gcc ttt att ctt gat ctc ccc      624
Ser Ser Leu Ile Gly Gly Leu Ile Asn Ala Phe Ile Leu Asp Leu Pro
    195                 200                 205

atc aat acc gcg ctg gca atg gcc tcc ggt ttc ggc tgg tat tct ctt      672
Ile Asn Thr Ala Leu Ala Met Ala Ser Gly Phe Gly Trp Tyr Ser Leu
    210                 215                 220

tcc ggt att tta ttg acc gaa tct ttt ggt ccg gta atc ggg agc gcg      720
Ser Gly Ile Leu Leu Thr Glu Ser Phe Gly Pro Val Ile Gly Ser Ala
225                 230                 235                 240

gcg ttt ttt aat gat ctg gcc cgt gaa ctg att gct att atg ttg atc      768
Ala Phe Phe Asn Asp Leu Ala Arg Glu Leu Ile Ala Ile Met Leu Ile
                245                 250                 255

cct ggg ctg att cgc cgc agc cgc tct act gca ctg ggc tta tgc ggt      816
Pro Gly Leu Ile Arg Arg Ser Arg Ser Thr Ala Leu Gly Leu Cys Gly
            260                 265                 270

gcc aca tca atg gat ttc acc ctg ccc gtt ctt caa cgt act ggc ggg      864
Ala Thr Ser Met Asp Phe Thr Leu Pro Val Leu Gln Arg Thr Gly Gly
        275                 280                 285

ctg gat atg gtc ccg gcg gca att gtt cac ggt ttt att ctt agc ctg      912
Leu Asp Met Val Pro Ala Ala Ile Val His Gly Phe Ile Leu Ser Leu
    290                 295                 300

tta gtg ccg atc ctc atc gcc ttt ttc tct gcg taa                      948
Leu Val Pro Ile Leu Ile Ala Phe Phe Ser Ala  *
305                 310                 315
```

<210> 169
<211> 291
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(291)

<400> 169

```
atg caa aac aca act cat gac aac gta att ctg gag ctc acc gtt cgc      48
Met Gln Asn Thr Thr His Asp Asn Val Ile Leu Glu Leu Thr Val Arg
 1               5                  10                  15

aac cat ccg ggc gta atg acc cac gtt tgt ggc ctt ttt gcc cgc cgc      96
Asn His Pro Gly Val Met Thr His Val Cys Gly Leu Phe Ala Arg Arg
            20                  25                  30

gct ttt aac gtt gaa ggc att ctt tgt ctg ccg att cag gac agc gac      144
Ala Phe Asn Val Glu Gly Ile Leu Cys Leu Pro Ile Gln Asp Ser Asp
        35                  40                  45

aaa agc cat atc tgg cta ctg gtc aat gac gac cag cgt ctg gag cag      192
Lys Ser His Ile Trp Leu Leu Val Asn Asp Asp Gln Arg Leu Glu Gln
    50                  55                  60
```

atg ata agc caa atc gat aag ctg gaa gat gtc gtg aaa gtg cag cgt 240
Met Ile Ser Gln Ile Asp Lys Leu Glu Asp Val Val Lys Val Gln Arg
65 70 75 80

aat cag tcc gat ccg acg atg ttt aac aag atc gcg gtg ttt ttt cag 288
Asn Gln Ser Asp Pro Thr Met Phe Asn Lys Ile Ala Val Phe Phe Gln
85 90 95

taa 291
*

<210> 170
<211> 1689
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1689)

<400> 170
atg gca agt tcg ggc aca aca tcg acg cgt aag cgc ttt acc ggc gca 48
Met Ala Ser Ser Gly Thr Thr Ser Thr Arg Lys Arg Phe Thr Gly Ala
1 5 10 15

gaa ttt atc gtt cat ttc ctg gaa cag cag ggc att aag att gtg aca 96
Glu Phe Ile Val His Phe Leu Glu Gln Gln Gly Ile Lys Ile Val Thr
20 25 30

ggc att ccg ggc ggt tct atc ctg cct gtt tac gat gcc tta agc caa 144
Gly Ile Pro Gly Gly Ser Ile Leu Pro Val Tyr Asp Ala Leu Ser Gln
35 40 45

agc acg caa atc cgc cat att ctg gcc cgt cat gaa cag ggc gcg ggc 192
Ser Thr Gln Ile Arg His Ile Leu Ala Arg His Glu Gln Gly Ala Gly
50 55 60

ttt atc gct cag gga atg gcg cgc acc gac ggt aaa ccg gcg gtc tgt 240
Phe Ile Ala Gln Gly Met Ala Arg Thr Asp Gly Lys Pro Ala Val Cys
65 70 75 80

atg gcc tgt agc gga ccg ggt gcg act aac ctg gtg acc gcc att gcc 288
Met Ala Cys Ser Gly Pro Gly Ala Thr Asn Leu Val Thr Ala Ile Ala
85 90 95

gat gcg cgg ctg gac tcc atc ccg ctg att tgc atc act ggt cag gtt 336
Asp Ala Arg Leu Asp Ser Ile Pro Leu Ile Cys Ile Thr Gly Gln Val
100 105 110

ccc gcc tcg atg atc ggc acc gac gcc ttc cag gaa gtg gac acc tac 384
Pro Ala Ser Met Ile Gly Thr Asp Ala Phe Gln Glu Val Asp Thr Tyr
115 120 125

ggc atc tct atc ccc atc acc aaa cac aac tat ctg gtc aga cat atc 432
Gly Ile Ser Ile Pro Ile Thr Lys His Asn Tyr Leu Val Arg His Ile
130 135 140

gaa gaa ctc ccg cag gtc atg agc gat gcc ttc cgc att gcg caa tca 480

Glu Glu Leu Pro Gln Val Met Ser Asp Ala Phe Arg Ile Ala Gln Ser
145 150 155 160

ggc cgc cca ggc ccg gtg tgg ata gac att cct aag gat gtg caa acg 528
Gly Arg Pro Gly Pro Val Trp Ile Asp Ile Pro Lys Asp Val Gln Thr
165 170 175

gca gtt ttt gag att gaa aca cag ccc gct atg gca gaa aaa gcc gcc 576
Ala Val Phe Glu Ile Glu Thr Gln Pro Ala Met Ala Glu Lys Ala Ala
180 185 190

gcc ccc gcc ttt agc gaa gaa agc att cgt gac gca gcg gcg atg att 624
Ala Pro Ala Phe Ser Glu Glu Ser Ile Arg Asp Ala Ala Ala Met Ile
195 200 205

aac gct gcc aaa cgc ccg gtg ctt tat ctg ggc ggc ggt gtg atc aat 672
Asn Ala Ala Lys Arg Pro Val Leu Tyr Leu Gly Gly Gly Val Ile Asn
210 215 220

gcg ccc gca cgg gtg cgt gaa ctg gcg gag aaa gcg caa ctg cct acc 720
Ala Pro Ala Arg Val Arg Glu Leu Ala Glu Lys Ala Gln Leu Pro Thr
225 230 235 240

acc atg act tta atg gcg ctg ggc atg ttg cca aaa gcg cat ccg ttg 768
Thr Met Thr Leu Met Ala Leu Gly Met Leu Pro Lys Ala His Pro Leu
245 250 255

tcg ctg ggt atg ctg ggg atg cac ggc gtg cgc agc acc aac tat att 816
Ser Leu Gly Met Leu Gly Met His Gly Val Arg Ser Thr Asn Tyr Ile
260 265 270

ttg cag gag gcg gat ttg ttg ata gtg ctc ggt gcg cgt ttt gat gac 864
Leu Gln Glu Ala Asp Leu Leu Ile Val Leu Gly Ala Arg Phe Asp Asp
275 280 285

cgg gcg att ggc aaa acc gag cag ttc tgt ccg aat gcc aaa atc att 912
Arg Ala Ile Gly Lys Thr Glu Gln Phe Cys Pro Asn Ala Lys Ile Ile
290 295 300

cat gtc gat atc gac cgt gca gag ctg ggt aaa atc aag cag ccg cac 960
His Val Asp Ile Asp Arg Ala Glu Leu Gly Lys Ile Lys Gln Pro His
305 310 315 320

gtg gcg att cag gcg gat gtt gat gac gtg ctg gcg cag ttg atc ccg 1008
Val Ala Ile Gln Ala Asp Val Asp Asp Val Leu Ala Gln Leu Ile Pro
325 330 335

ctg gtg gaa gcg caa ccg cgt gca gag tgg cac cag ttg gta gcg gat 1056
Leu Val Glu Ala Gln Pro Arg Ala Glu Trp His Gln Leu Val Ala Asp
340 345 350

ttg cag cgt gag ttt ccg tgt cca atc ccg aaa gcg tgc gat ccg tta 1104
Leu Gln Arg Glu Phe Pro Cys Pro Ile Pro Lys Ala Cys Asp Pro Leu
355 360 365

agc cat tac ggc ctg atc aac gcc gtt gcc gcc tgt gtc gat gac aat 1152
Ser His Tyr Gly Leu Ile Asn Ala Val Ala Ala Cys Val Asp Asp Asn
370 375 380

gca att atc acc acc gac gtt ggt cag cat cag atg tgg acc gcg caa 1200
Ala Ile Ile Thr Thr Asp Val Gly Gln His Gln Met Trp Thr Ala Gln

385 390 395 400

gct tat ccg ctc aat cgc cca cgc cag tgg ctg acc tcc ggt ggg ctg 1248
Ala Tyr Pro Leu Asn Arg Pro Arg Gln Trp Leu Thr Ser Gly Gly Leu
405 410 415

ggc acg atg ggt ttt ggc ctg cct gcg gcg att ggc gct gcg ctg gcg 1296
Gly Thr Met Gly Phe Gly Leu Pro Ala Ala Ile Gly Ala Ala Leu Ala
420 425 430

aac ccg gat cgc aaa gtg ttg tgt ttc tcc ggc gac ggc agc ctg atg 1344
Asn Pro Asp Arg Lys Val Leu Cys Phe Ser Gly Asp Gly Ser Leu Met
435 440 445

atg aat att cag gag atg gcg acc gcc agt gaa aat cag ctg gat gtc 1392
Met Asn Ile Gln Glu Met Ala Thr Ala Ser Glu Asn Gln Leu Asp Val
450 455 460

aaa atc att ctg atg aac aac gaa gcg ctg ggg ctg gtg cat cag caa 1440
Lys Ile Ile Leu Met Asn Asn Glu Ala Leu Gly Leu Val His Gln Gln
465 470 475 480

cag agt ctg ttc tac gag caa ggc gtt ttt gcc gcc acc tat ccg ggc 1488
Gln Ser Leu Phe Tyr Glu Gln Gly Val Phe Ala Ala Thr Tyr Pro Gly
485 490 495

aaa atc aac ttt atg cag att gcc gcc gga ttc ggc ctc gaa acc tgt 1536
Lys Ile Asn Phe Met Gln Ile Ala Ala Gly Phe Gly Leu Glu Thr Cys
500 505 510

gat ttg aat aac gaa gcc gat ccg cag gct tca ttg cag gaa atc atc 1584
Asp Leu Asn Asn Glu Ala Asp Pro Gln Ala Ser Leu Gln Glu Ile Ile
515 520 525

aat cgc cct ggc ccg gcg ctg atc cat gtg cgc att gat gcc gaa gaa 1632
Asn Arg Pro Gly Pro Ala Leu Ile His Val Arg Ile Asp Ala Glu Glu
530 535 540

aaa gtt tac ccg atg gtg ccg cca ggt gcg gcg aat act gaa atg gtg 1680
Lys Val Tyr Pro Met Val Pro Pro Gly Ala Ala Asn Thr Glu Met Val
545 550 555 560

ggg gaa taa 1689
Gly Glu *

<210> 171
<211> 99
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(99)

<400> 171
atg act act tcc atg ctc aac gca aaa cta cta cca act gcg cca tcc 48
Met Thr Thr Ser Met Leu Asn Ala Lys Leu Leu Pro Thr Ala Pro Ser
1 5 10 15

gcc gca gtg gtc gtc gtg cgt gtg gtg gtg gtc gtc ggc aat gcg ccg 96
Ala Ala Val Val Val Val Arg Val Val Val Val Val Gly Asn Ala Pro
20 25 30

tag 99
*

<210> 172
<211> 1017
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1017)

<400> 172
atg ttt gtc atc tgg agc cat aga aca ggg ttc atc atg agt cat caa 48
Met Phe Val Ile Trp Ser His Arg Thr Gly Phe Ile Met Ser His Gln
1 5 10 15

ctt acc ttc gcc gac agt gaa ttc agc agt aag cgc cgt cag acc aga 96
Leu Thr Phe Ala Asp Ser Glu Phe Ser Ser Lys Arg Arg Gln Thr Arg
20 25 30

aaa gag att ttc ttg tcc cgc atg gag cag att ctg cca tgg caa aac 144
Lys Glu Ile Phe Leu Ser Arg Met Glu Gln Ile Leu Pro Trp Gln Asn
35 40 45

atg gtg gaa gtc atc gag ccg ttt tac ccc aag gct ggt aat ggc cgg 192
Met Val Glu Val Ile Glu Pro Phe Tyr Pro Lys Ala Gly Asn Gly Arg
50 55 60

cga cct tat ccg ctg gaa acc atg cta cgc att cac tgc atg cag cat 240
Arg Pro Tyr Pro Leu Glu Thr Met Leu Arg Ile His Cys Met Gln His
65 70 75 80

tgg tac aac ctg agc gat ggc gcg atg gaa gat gct ctg tac gaa atc 288
Trp Tyr Asn Leu Ser Asp Gly Ala Met Glu Asp Ala Leu Tyr Glu Ile
85 90 95

gcc tcc atg cgt ctg ttt gcc cgg tta tcc ctg gat agc gcc ttg ccg 336
Ala Ser Met Arg Leu Phe Ala Arg Leu Ser Leu Asp Ser Ala Leu Pro
100 105 110

gac cgc acc acc atc atg aat ttc cgc cac ctg ctg gag cag cat caa 384
Asp Arg Thr Thr Ile Met Asn Phe Arg His Leu Leu Glu Gln His Gln
115 120 125

ctg gcc cgc caa ttg ttc aag acc atc aat cgc tgg ctg gcc gaa gca 432
Leu Ala Arg Gln Leu Phe Lys Thr Ile Asn Arg Trp Leu Ala Glu Ala
130 135 140

ggc gtc atg atg act caa ggc acc ttg gtc gat gcc acc atc att gag 480
Gly Val Met Met Thr Gln Gly Thr Leu Val Asp Ala Thr Ile Ile Glu
145 150 155 160

gca ccc agc tcg acc aag aac aaa gag cag caa cgc gat ccg gag atg 528
Ala Pro Ser Ser Thr Lys Asn Lys Glu Gln Gln Arg Asp Pro Glu Met

```
                165                 170                 175

cat cag acc aag aaa ggc aat cag tgg cac ttt ggc atg aag gcc cac      576
His Gln Thr Lys Lys Gly Asn Gln Trp His Phe Gly Met Lys Ala His
            180                 185                 190

att ggt gtc gat gcc aag agt ggc ctg acc cac agc ctg gtc acc acc      624
Ile Gly Val Asp Ala Lys Ser Gly Leu Thr His Ser Leu Val Thr Thr
        195                 200                 205

gcg gcc aac gag cat gac ctc aat cag ctg ggt aat ctg ctg cat gga      672
Ala Ala Asn Glu His Asp Leu Asn Gln Leu Gly Asn Leu Leu His Gly
    210                 215                 220

gag gag caa ttt gtc tca gcc gat gcc ggc tac caa ggg gcg cca cag      720
Glu Glu Gln Phe Val Ser Ala Asp Ala Gly Tyr Gln Gly Ala Pro Gln
225                 230                 235                 240

cgc gag gag ctg gcc gag gtg gat gtg gac tgg ctg atc gcc gag cgc      768
Arg Glu Glu Leu Ala Glu Val Asp Val Asp Trp Leu Ile Ala Glu Arg
                245                 250                 255

ccc ggc aag gta aga acc ttg aaa cag cat cca cgc aag aac aaa acg      816
Pro Gly Lys Val Arg Thr Leu Lys Gln His Pro Arg Lys Asn Lys Thr
            260                 265                 270

gcc atc aac atc gaa tac atg aaa gcc agc atc cgg gcc agg gtg gag      864
Ala Ile Asn Ile Glu Tyr Met Lys Ala Ser Ile Arg Ala Arg Val Glu
        275                 280                 285

cac cca ttt cgc atc atc aag cga cag ttc ggc ttc gtg aaa gcc aga      912
His Pro Phe Arg Ile Ile Lys Arg Gln Phe Gly Phe Val Lys Ala Arg
    290                 295                 300

tac aag ggg ttg ctg aaa aac gat aac caa ctg gcg atg tta ttc acg      960
Tyr Lys Gly Leu Leu Lys Asn Asp Asn Gln Leu Ala Met Leu Phe Thr
305                 310                 315                 320

ctg gcc aac ctg ttt cgg gcg gac caa atg ata cgt cag tgg gag aga      1008
Leu Ala Asn Leu Phe Arg Ala Asp Gln Met Ile Arg Gln Trp Glu Arg
                325                 330                 335

tct cac taa                                                          1017
Ser His  *
```

<210> 173
<211> 474
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(474)

<400> 173

```
atg gta tat ata ata atc gtt tcc cac gga cat gaa gac tac atc aaa      48
Met Val Tyr Ile Ile Ile Val Ser His Gly His Glu Asp Tyr Ile Lys
 1               5                  10                  15
```

aaa tta ctc gaa aat ctt aat gct gac gat gag cac tac aag att atc 96
Lys Leu Leu Glu Asn Leu Asn Ala Asp Asp Glu His Tyr Lys Ile Ile
20 25 30

gta cgc gac aac aaa gac tct cta tta ttg aaa caa ata tgc cag cat 144
Val Arg Asp Asn Lys Asp Ser Leu Leu Leu Lys Gln Ile Cys Gln His
35 40 45

tat gca ggc ctg gac tat att agt gga ggt gta tac ggc ttt ggt cat 192
Tyr Ala Gly Leu Asp Tyr Ile Ser Gly Gly Val Tyr Gly Phe Gly His
50 55 60

aat aat aat att gcg gtg gcg tat gta aag gaa aaa tat aga ccc gca 240
Asn Asn Asn Ile Ala Val Ala Tyr Val Lys Glu Lys Tyr Arg Pro Ala
65 70 75 80

gat gat gat tac att ttg ttt ttg aat ccc gat atc atc atg aag cat 288
Asp Asp Asp Tyr Ile Leu Phe Leu Asn Pro Asp Ile Ile Met Lys His
85 90 95

gat gat ttg ctg aca tat att aaa tat gtc gaa agt aag cgt tat gct 336
Asp Asp Leu Leu Thr Tyr Ile Lys Tyr Val Glu Ser Lys Arg Tyr Ala
100 105 110

ttt agt aca tta tgc ctg ttc cga gat gaa gcg aaa tct tta cat gat 384
Phe Ser Thr Leu Cys Leu Phe Arg Asp Glu Ala Lys Ser Leu His Asp
115 120 125

tat tcc gta aga aaa ttt cct gtg ctt tct gat ttt att gtg tca ttt 432
Tyr Ser Val Arg Lys Phe Pro Val Leu Ser Asp Phe Ile Val Ser Phe
130 135 140

atg tta ggg att aag gaa ggt gcg aac aag tcc ctg ata tga 474
Met Leu Gly Ile Lys Glu Gly Ala Asn Lys Ser Leu Ile *
145 150 155

<210> 174
<211> 1119
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1119)

<400> 174
atg gga aaa agc ata gtc gtt gtt tct gcg gtc aat ttt acc act ggc 48
Met Gly Lys Ser Ile Val Val Val Ser Ala Val Asn Phe Thr Thr Gly
1 5 10 15

ggt cca ttt acc att ttg aaa aaa ttt ttg gca gca act aat aat aaa 96
Gly Pro Phe Thr Ile Leu Lys Lys Phe Leu Ala Ala Thr Asn Asn Lys
20 25 30

gaa aat gtc agt ttt atc gca tta gtc cat tct gct aaa gag tta aaa 144
Glu Asn Val Ser Phe Ile Ala Leu Val His Ser Ala Lys Glu Leu Lys
35 40 45

gaa agt tat cca tgg gtt aaa ttc att gag ttt cct gag gtt aaa ggg 192

Glu Ser Tyr Pro Trp Val Lys Phe Ile Glu Phe Pro Glu Val Lys Gly
50 55 60

tcg tgg cta aaa cgt ttg cac ttt gaa tat gta gtt tgt aaa aaa ctt 240
Ser Trp Leu Lys Arg Leu His Phe Glu Tyr Val Val Cys Lys Lys Leu
65 70 75 80

tca aaa gag ctg aat gct acg cat tgg att tgt ctg cat gat att acg 288
Ser Lys Glu Leu Asn Ala Thr His Trp Ile Cys Leu His Asp Ile Thr
85 90 95

gcc aat gtc gtc act aaa aaa aga tat gtg tat tgt cat aac cct gca 336
Ala Asn Val Val Thr Lys Lys Arg Tyr Val Tyr Cys His Asn Pro Ala
100 105 110

cct ttt tat aaa gga att tta ttc cgt gaa att ctt atg gag cct agc 384
Pro Phe Tyr Lys Gly Ile Leu Phe Arg Glu Ile Leu Met Glu Pro Ser
115 120 125

ttt ttc tta ttt aaa atg cta tac ggg ctg ata tat aaa ata aac att 432
Phe Phe Leu Phe Lys Met Leu Tyr Gly Leu Ile Tyr Lys Ile Asn Ile
130 135 140

aaa aaa aat act gca gtg ttt gtt caa caa ttc tgg atg aaa gaa aaa 480
Lys Lys Asn Thr Ala Val Phe Val Gln Gln Phe Trp Met Lys Glu Lys
145 150 155 160

ttt atc aag aaa tat tct ata aat aac atc att gtc agt cgg cca gaa 528
Phe Ile Lys Lys Tyr Ser Ile Asn Asn Ile Ile Val Ser Arg Pro Glu
165 170 175

att aaa tta tct gat aaa agc caa ctt act gat gat gat tct caa ttt 576
Ile Lys Leu Ser Asp Lys Ser Gln Leu Thr Asp Asp Asp Ser Gln Phe
180 185 190

aag aat aac cct tct gag ttg aca ata ttt tac cct gct gtt cca cga 624
Lys Asn Asn Pro Ser Glu Leu Thr Ile Phe Tyr Pro Ala Val Pro Arg
195 200 205

gta ttt aaa aat tac gag ctt att att agt gca gca agg aaa ttg aaa 672
Val Phe Lys Asn Tyr Glu Leu Ile Ile Ser Ala Ala Arg Lys Leu Lys
210 215 220

gaa caa tcc aat att aaa ttt ctg ctt act atc agt ggt aca gaa aat 720
Glu Gln Ser Asn Ile Lys Phe Leu Leu Thr Ile Ser Gly Thr Glu Asn
225 230 235 240

gcg tat gca aaa tat att atc agt ctt gca gaa gga ctg gat aat gtt 768
Ala Tyr Ala Lys Tyr Ile Ile Ser Leu Ala Glu Gly Leu Asp Asn Val
245 250 255

cat ttc ctc ggg tac ttg gat aaa gaa aaa atc gat cat tgt tat aat 816
His Phe Leu Gly Tyr Leu Asp Lys Glu Lys Ile Asp His Cys Tyr Asn
260 265 270

att tca gat ata gtt tgt ttt ccc tct agg tta gaa aca tgg gga ttg 864
Ile Ser Asp Ile Val Cys Phe Pro Ser Arg Leu Glu Thr Trp Gly Leu
275 280 285

ccg ttg tct gaa gct aaa gag cga ggt aag tgg gta tta gca tca gat 912
Pro Leu Ser Glu Ala Lys Glu Arg Gly Lys Trp Val Leu Ala Ser Asp

290 295 300

ttc cca ttt act aga gaa act ctt ggt agt tat gaa aag aaa gct ttt 960
Phe Pro Phe Thr Arg Glu Thr Leu Gly Ser Tyr Glu Lys Lys Ala Phe
305 310 315 320

ttt gat tct aat aac gat gac atg tta gtt aaa ctt att att gac ttc 1008
Phe Asp Ser Asn Asn Asp Asp Met Leu Val Lys Leu Ile Ile Asp Phe
325 330 335

aaa aaa ggt aac ctc aaa aaa gat atc tct gat gca aat ttc att tat 1056
Lys Lys Gly Asn Leu Lys Lys Asp Ile Ser Asp Ala Asn Phe Ile Tyr
340 345 350

cgt aat gaa aat gta tta gtt ggg ttt gat gaa cta gtt aat ttt att 1104
Arg Asn Glu Asn Val Leu Val Gly Phe Asp Glu Leu Val Asn Phe Ile
355 360 365

act gaa gaa cat tga 1119
Thr Glu Glu His *
370

<210> 175
<211> 591
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(591)

<400> 175
atg atc tta aaa ctc gct aaa cga tat ggt ctc tgt ggt ttt att cgg 48
Met Ile Leu Lys Leu Ala Lys Arg Tyr Gly Leu Cys Gly Phe Ile Arg
1 5 10 15

ctt gtt aga gat gtc tta ttg act cgt gta ttt tac cgg aac tgt aga 96
Leu Val Arg Asp Val Leu Leu Thr Arg Val Phe Tyr Arg Asn Cys Arg
20 25 30

att att cga ttt ccc tgc tat att cgc aat gat ggt agc att aat ttt 144
Ile Ile Arg Phe Pro Cys Tyr Ile Arg Asn Asp Gly Ser Ile Asn Phe
35 40 45

ggt gaa aat ttc aca agt gga gtc ggt ctc agg ctg gat gca ttt gga 192
Gly Glu Asn Phe Thr Ser Gly Val Gly Leu Arg Leu Asp Ala Phe Gly
50 55 60

cgt ggc gtg att ttt ttt tcc gat aat gtg caa gtt aac gac tat gtt 240
Arg Gly Val Ile Phe Phe Ser Asp Asn Val Gln Val Asn Asp Tyr Val
65 70 75 80

cat atc gcc tca att gag agc gtt acg ata ggt cgg gat acg ctt att 288
His Ile Ala Ser Ile Glu Ser Val Thr Ile Gly Arg Asp Thr Leu Ile
85 90 95

gca agt aaa gta ttt att acc gat cat aat cac ggt tcc ttt aag cac 336
Ala Ser Lys Val Phe Ile Thr Asp His Asn His Gly Ser Phe Lys His
100 105 110

```
tct gat cca atg agt tcg cca aat ata cct cca gac atg cgc acg ttg    384
Ser Asp Pro Met Ser Ser Pro Asn Ile Pro Pro Asp Met Arg Thr Leu
        115                 120                 125

gaa tct tca gct gtt gta att ggc cag agg gtt tgg ttg ggt gag aat    432
Glu Ser Ser Ala Val Val Ile Gly Gln Arg Val Trp Leu Gly Glu Asn
    130                 135                 140

gtg acg gtt ttg cct gga aca att att ggt aat gga gtc gta gtc ggc    480
Val Thr Val Leu Pro Gly Thr Ile Ile Gly Asn Gly Val Val Val Gly
145                 150                 155                 160

gcc aat tct gtt gtt aga ggt tct att ccc gaa aat act gtc att gcg    528
Ala Asn Ser Val Val Arg Gly Ser Ile Pro Glu Asn Thr Val Ile Ala
                165                 170                 175

gga gta cca gca aaa atc ata aag aaa tac aat cat gag acc aaa tta    576
Gly Val Pro Ala Lys Ile Ile Lys Lys Tyr Asn His Glu Thr Lys Leu
            180                 185                 190

tgg gaa aaa gca tag                                                591
Trp Glu Lys Ala  *
        195
```

<210> 176
<211> 993
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(993)

<400> 176

```
atg tat ttt ttg aat gat tta aat ttc tct aga cgc gat gct gga ttt    48
Met Tyr Phe Leu Asn Asp Leu Asn Phe Ser Arg Arg Asp Ala Gly Phe
 1               5                  10                  15

aaa gca aga aaa gat gca ctg gac att gct tca gat tat gaa aac att    96
Lys Ala Arg Lys Asp Ala Leu Asp Ile Ala Ser Asp Tyr Glu Asn Ile
            20                  25                  30

tct gtt gtt aac att cct cta tgg ggt gga gta gtc cag aga att att    144
Ser Val Val Asn Ile Pro Leu Trp Gly Gly Val Val Gln Arg Ile Ile
        35                  40                  45

agt tct gtt aag ctt agt aca ttt ctc tgc ggt ctt gaa aat aaa gat    192
Ser Ser Val Lys Leu Ser Thr Phe Leu Cys Gly Leu Glu Asn Lys Asp
    50                  55                  60

gtt tta att ttc aat ttc ccg atg gcc aaa cca ttt tgg cat ata ttg    240
Val Leu Ile Phe Asn Phe Pro Met Ala Lys Pro Phe Trp His Ile Leu
65                  70                  75                  80

tca ttc ttt cac cgc ctt cta aaa ttt aga ata gta cct ctg att cat    288
Ser Phe Phe His Arg Leu Leu Lys Phe Arg Ile Val Pro Leu Ile His
                85                  90                  95

gat att gat gaa tta aga gga gga ggg ggt agt gat tct gtg cgg ctt    336
Asp Ile Asp Glu Leu Arg Gly Gly Gly Gly Ser Asp Ser Val Arg Leu
```

```
            100                 105                 110

gct acc tgt gat atg gtc ata agt cac aat cca caa atg aca aag tac     384
Ala Thr Cys Asp Met Val Ile Ser His Asn Pro Gln Met Thr Lys Tyr
        115                 120                 125

ctt agt aaa tat atg tct cag gat aaa atc aaa gac ata aaa ata ttt     432
Leu Ser Lys Tyr Met Ser Gln Asp Lys Ile Lys Asp Ile Lys Ile Phe
    130                 135                 140

gat tac ctc gtc tca tct gat gtg gag cat cga gat gtt acg gat aag     480
Asp Tyr Leu Val Ser Ser Asp Val Glu His Arg Asp Val Thr Asp Lys
145                 150                 155                 160

caa cga ggg gtc ata tat gct ggc aac ctt tct agg cat aaa tgt tct     528
Gln Arg Gly Val Ile Tyr Ala Gly Asn Leu Ser Arg His Lys Cys Ser
                165                 170                 175

ttc ata tat act gaa gga tgc gat ttt act ctc ttt ggt gtc aac tat     576
Phe Ile Tyr Thr Glu Gly Cys Asp Phe Thr Leu Phe Gly Val Asn Tyr
            180                 185                 190

gaa aat aaa gat aat cct aaa tat ctt gga agt ttt gat gct caa tct     624
Glu Asn Lys Asp Asn Pro Lys Tyr Leu Gly Ser Phe Asp Ala Gln Ser
        195                 200                 205

ccg gaa aag att aac ctc cca ggc atg caa ttt gga ctc att tgg gat     672
Pro Glu Lys Ile Asn Leu Pro Gly Met Gln Phe Gly Leu Ile Trp Asp
    210                 215                 220

gga gat tct gtc gaa acc tgt agt ggt gcc ttt ggc gac tat tta aag     720
Gly Asp Ser Val Glu Thr Cys Ser Gly Ala Phe Gly Asp Tyr Leu Lys
225                 230                 235                 240

ttt aat aac cct cat aag aca tct ctt tat ctt tca atg gaa ctt cca     768
Phe Asn Asn Pro His Lys Thr Ser Leu Tyr Leu Ser Met Glu Leu Pro
                245                 250                 255

gta ttt ata tgg gat aaa gcc gcc ctt gcg gat ttc att gta gat aat     816
Val Phe Ile Trp Asp Lys Ala Ala Leu Ala Asp Phe Ile Val Asp Asn
            260                 265                 270

aga ata gga tat gca gtg gga tca atc aaa gaa atg caa gag att gtt     864
Arg Ile Gly Tyr Ala Val Gly Ser Ile Lys Glu Met Gln Glu Ile Val
        275                 280                 285

gac tcc atg aca ata gaa act tat aag caa att agt gag aat aca aaa     912
Asp Ser Met Thr Ile Glu Thr Tyr Lys Gln Ile Ser Glu Asn Thr Lys
    290                 295                 300

att att tct cag aaa att cga aca gga agt tac ttc agg gat gtt ctt     960
Ile Ile Ser Gln Lys Ile Arg Thr Gly Ser Tyr Phe Arg Asp Val Leu
305                 310                 315                 320

gaa gag gtg atc gat gat ctt aaa act cgc taa                         993
Glu Glu Val Ile Asp Asp Leu Lys Thr Arg  *
                325                 330
```

<210> 177
<211> 1167

<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1167)

<400> 177
atg atc tat ctt gta att agt gtc ttt ctc att aca gca ttt atc tgt 48
Met Ile Tyr Leu Val Ile Ser Val Phe Leu Ile Thr Ala Phe Ile Cys
1 5 10 15

tta tat ctt aag aag gat ata ttt tat cca gcc gta tgc gtt aat atc 96
Leu Tyr Leu Lys Lys Asp Ile Phe Tyr Pro Ala Val Cys Val Asn Ile
20 25 30

atc ttc gca ctg gtc tta ttg gga tat gaa ata acg tca gat ata tat 144
Ile Phe Ala Leu Val Leu Leu Gly Tyr Glu Ile Thr Ser Asp Ile Tyr
35 40 45

gct ttt cag tta aat gac gct acg ttg att ttt cta ctt tgc aat gtt 192
Ala Phe Gln Leu Asn Asp Ala Thr Leu Ile Phe Leu Leu Cys Asn Val
50 55 60

ttg aca ttt acc ctg tca tgt tta ttg acg gaa agt gta tta gat cta 240
Leu Thr Phe Thr Leu Ser Cys Leu Leu Thr Glu Ser Val Leu Asp Leu
65 70 75 80

aat atc aga aaa gtc aat aat gct att tat agc ata cca tcg aag aaa 288
Asn Ile Arg Lys Val Asn Asn Ala Ile Tyr Ser Ile Pro Ser Lys Lys
85 90 95

gtg cat aat gta ggc ttg tta gtt att tct ttt tcg atg ata tat ata 336
Val His Asn Val Gly Leu Leu Val Ile Ser Phe Ser Met Ile Tyr Ile
100 105 110

tgc atg agg tta agt aac tac cag ttc ggg act agc tta ctt agc tat 384
Cys Met Arg Leu Ser Asn Tyr Gln Phe Gly Thr Ser Leu Leu Ser Tyr
115 120 125

atg aat ttg ata aga gat gct gat gtt gaa gac aca tca aga aat ttc 432
Met Asn Leu Ile Arg Asp Ala Asp Val Glu Asp Thr Ser Arg Asn Phe
130 135 140

tca gca tac atg cag cca atc att cta act act ttt gct tta ttt att 480
Ser Ala Tyr Met Gln Pro Ile Ile Leu Thr Thr Phe Ala Leu Phe Ile
145 150 155 160

tgg tct aaa aaa ttt act aat aca aag gta agt aaa aca ttt act tta 528
Trp Ser Lys Lys Phe Thr Asn Thr Lys Val Ser Lys Thr Phe Thr Leu
165 170 175

ctt gtt ttt att gta ttc atc ttt gca att ata ctg aat act ggt aag 576
Leu Val Phe Ile Val Phe Ile Phe Ala Ile Ile Leu Asn Thr Gly Lys
180 185 190

caa att gtc ttt atg gtt atc atc tct tat gca ttc atc gta ggt gtt 624
Gln Ile Val Phe Met Val Ile Ile Ser Tyr Ala Phe Ile Val Gly Val
195 200 205

aat aga gta aaa cat tat gtt tat ctt att aca gct gta ggt gtt cta 672

```
Asn Arg Val Lys His Tyr Val Tyr Leu Ile Thr Ala Val Gly Val Leu
    210                 215                 220

ttc tcc ttg tat atg ctc ttt tta cgt gga ctg cct ggg ggg atg gca     720
Phe Ser Leu Tyr Met Leu Phe Leu Arg Gly Leu Pro Gly Gly Met Ala
225                 230                 235                 240

tat tat cta tcc atg tat ttg gtc agc cct ata atc gcg ttt cag gag     768
Tyr Tyr Leu Ser Met Tyr Leu Val Ser Pro Ile Ile Ala Phe Gln Glu
                245                 250                 255

ttt tat ttt cag caa gta tct aac tct gcc agt tct cat gtc ttt tgg     816
Phe Tyr Phe Gln Gln Val Ser Asn Ser Ala Ser Ser His Val Phe Trp
            260                 265                 270

ttt ttt gaa agg ctg atg ggg cta tta aca ggt gga gtc tct atg tcg     864
Phe Phe Glu Arg Leu Met Gly Leu Leu Thr Gly Gly Val Ser Met Ser
        275                 280                 285

ttg cat aaa gaa ttt gtg tgg gtg ggt ttg cca aca aat gtt tat act     912
Leu His Lys Glu Phe Val Trp Val Gly Leu Pro Thr Asn Val Tyr Thr
    290                 295                 300

gct ttt tcg gat tat gtt tat att tcc gcg gag cta agc tat ttg atg     960
Ala Phe Ser Asp Tyr Val Tyr Ile Ser Ala Glu Leu Ser Tyr Leu Met
305                 310                 315                 320

atg gtt att cat ggc tgt att tca ggt gtt tta tgg aga ttg tct cga     1008
Met Val Ile His Gly Cys Ile Ser Gly Val Leu Trp Arg Leu Ser Arg
                325                 330                 335

aat tac ata tct gtg aaa ata ttt tat tca tat ttt att tat acc ttt     1056
Asn Tyr Ile Ser Val Lys Ile Phe Tyr Ser Tyr Phe Ile Tyr Thr Phe
            340                 345                 350

tct ttc att ttt tat cat gaa agc ttc atg act aat att agc agt tgg     1104
Ser Phe Ile Phe Tyr His Glu Ser Phe Met Thr Asn Ile Ser Ser Trp
        355                 360                 365

ata caa ata act ctt tgt atc ata gta ttc tct caa ttt ctt aag gcc     1152
Ile Gln Ile Thr Leu Cys Ile Ile Val Phe Ser Gln Phe Leu Lys Ala
    370                 375                 380

cag aaa ata aag tga                                                 1167
Gln Lys Ile Lys  *
385
```

<210> 178
<211> 1104
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1104)

<400> 178

```
atg tac gat tat atc att gtt ggt tct ggt ttg ttt ggt gcc gtt tgt     48
Met Tyr Asp Tyr Ile Ile Val Gly Ser Gly Leu Phe Gly Ala Val Cys
 1               5                  10                  15
```

```
gcg aat gag tta aaa aag cta aac aaa aaa gtt tta gtg att gag aaa     96
Ala Asn Glu Leu Lys Lys Leu Asn Lys Lys Val Leu Val Ile Glu Lys
            20                  25                  30

aga aat cat atc ggt gga aat gcg tac aca gag gac tgt gag ggt atc    144
Arg Asn His Ile Gly Gly Asn Ala Tyr Thr Glu Asp Cys Glu Gly Ile
        35                  40                  45

cag att cat aaa tat ggt gca cat att ttt cat acc aat gat aaa tat    192
Gln Ile His Lys Tyr Gly Ala His Ile Phe His Thr Asn Asp Lys Tyr
    50                  55                  60

ata tgg gat tac gtt aat gat tta gta gaa ttt aat cgt ttt act aat    240
Ile Trp Asp Tyr Val Asn Asp Leu Val Glu Phe Asn Arg Phe Thr Asn
65                  70                  75                  80

tct cca ctg gcg att tat aaa gac aaa tta ttc aac ctt cct ttt aat    288
Ser Pro Leu Ala Ile Tyr Lys Asp Lys Leu Phe Asn Leu Pro Phe Asn
                85                  90                  95

atg aat act ttc cac caa atg tgg gga gtt aaa gat cct caa gaa gct    336
Met Asn Thr Phe His Gln Met Trp Gly Val Lys Asp Pro Gln Glu Ala
            100                 105                 110

caa aat atc att aat gct cag aaa aaa aag tat ggt gac aag gta cct    384
Gln Asn Ile Ile Asn Ala Gln Lys Lys Lys Tyr Gly Asp Lys Val Pro
        115                 120                 125

gaa aat ttg gag gag cag gcg att tca tta gtt ggg gag gac tta tac    432
Glu Asn Leu Glu Glu Gln Ala Ile Ser Leu Val Gly Glu Asp Leu Tyr
    130                 135                 140

caa gca ttg ata aag ggt tat acg gag aag cag tgg gga aga agt gca    480
Gln Ala Leu Ile Lys Gly Tyr Thr Glu Lys Gln Trp Gly Arg Ser Ala
145                 150                 155                 160

aaa gaa ttg cct gca ttt att att aag cga atc cca gtg aga ttt acg    528
Lys Glu Leu Pro Ala Phe Ile Ile Lys Arg Ile Pro Val Arg Phe Thr
                165                 170                 175

ttt gat aac aat tat ttt tcc gat cgc tat caa ggt att ccg gtg gga    576
Phe Asp Asn Asn Tyr Phe Ser Asp Arg Tyr Gln Gly Ile Pro Val Gly
            180                 185                 190

ggc tac act aag ctt att gaa aaa atg ctt gaa ggt gtg gac gta aaa    624
Gly Tyr Thr Lys Leu Ile Glu Lys Met Leu Glu Gly Val Asp Val Lys
        195                 200                 205

tta ggc att gat ttt ttg aaa gac aaa gat tct cta gcg agt aaa gcc    672
Leu Gly Ile Asp Phe Leu Lys Asp Lys Asp Ser Leu Ala Ser Lys Ala
    210                 215                 220

cat aga atc atc tac act gga ccc att gat cag tac ttc gac tat agg    720
His Arg Ile Ile Tyr Thr Gly Pro Ile Asp Gln Tyr Phe Asp Tyr Arg
225                 230                 235                 240

ttt gga gcg tta gaa tat cgc tct tta aaa ttt gag acg gaa cgc cat    768
Phe Gly Ala Leu Glu Tyr Arg Ser Leu Lys Phe Glu Thr Glu Arg His
                245                 250                 255
```

gaa ttt cca aac ttc caa ggg aat gca gta ata aat ttc act gat gct 816
Glu Phe Pro Asn Phe Gln Gly Asn Ala Val Ile Asn Phe Thr Asp Ala
260 265 270

aat gta cca tat acc aga ata att gag cat aaa cat ttt gac tat gtt 864
Asn Val Pro Tyr Thr Arg Ile Ile Glu His Lys His Phe Asp Tyr Val
275 280 285

gag aca aag cat acg gtt gtt aca aaa gaa tat cca tta gag tgg aaa 912
Glu Thr Lys His Thr Val Val Thr Lys Glu Tyr Pro Leu Glu Trp Lys
290 295 300

gtt ggc gac gaa ccc tac tat cca gtt aat gat aat aaa aac atg gag 960
Val Gly Asp Glu Pro Tyr Tyr Pro Val Asn Asp Asn Lys Asn Met Glu
305 310 315 320

ctt ttt aag aaa tat aga gag tta gct agc aga gaa gac aag gtt ata 1008
Leu Phe Lys Lys Tyr Arg Glu Leu Ala Ser Arg Glu Asp Lys Val Ile
325 330 335

ttt ggc ggg cgt ttg gcc gag tat aaa tat tat gat atg cat caa gtg 1056
Phe Gly Gly Arg Leu Ala Glu Tyr Lys Tyr Tyr Asp Met His Gln Val
340 345 350

ata tct gcc gct ctt tat caa gtg aaa aat ata atg agt acg gat taa 1104
Ile Ser Ala Ala Leu Tyr Gln Val Lys Asn Ile Met Ser Thr Asp *
355 360 365

<210> 179
<211> 693
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(693)

<400> 179
ttg aca tca tca ata acc aat gaa att atg caa tta tat acg gat agg 48
Met Thr Ser Ser Ile Thr Asn Glu Ile Met Gln Leu Tyr Thr Asp Arg
1 5 10 15

gag gtt ctt aac atg ggg ctt tgt agt cgt tat aaa agt ctt aca tgc 96
Glu Val Leu Asn Met Gly Leu Cys Ser Arg Tyr Lys Ser Leu Thr Cys
20 25 30

aat agt tgc tca atg cat tgc caa ata atg cca gaa gag tca ccg cgt 144
Asn Ser Cys Ser Met His Cys Gln Ile Met Pro Glu Glu Ser Pro Arg
35 40 45

ttg cag tat tgc gca aac tcg tgt ttt tgt atg tgg ccc gaa gaa agt 192
Leu Gln Tyr Cys Ala Asn Ser Cys Phe Cys Met Trp Pro Glu Glu Ser
50 55 60

tca tat ttt aat cgt ggt gtg gta gaa ggt atc tta aca aaa aac cac 240
Ser Tyr Phe Asn Arg Gly Val Val Glu Gly Ile Leu Thr Lys Asn His
65 70 75 80

aat gcg aga tta agc gga tat att ttt gtc gat ttt tca gta agt ttt 288

```
Asn Ala Arg Leu Ser Gly Tyr Ile Phe Val Asp Phe Ser Val Ser Phe
                85                  90                  95

tta cgt cta ttc ctt gaa aaa gac tgg att gac tat ctt gca agt act    336
Leu Arg Leu Phe Leu Glu Lys Asp Trp Ile Asp Tyr Leu Ala Ser Thr
            100                 105                 110

gac atg ggg att gtt tta gtc agc gac cgc aat atg caa tcg tta gct    384
Asp Met Gly Ile Val Leu Val Ser Asp Arg Asn Met Gln Ser Leu Ala
        115                 120                 125

aat tac tgg cga aaa cac aac tca gca ata tct gct gtc ata tat aat    432
Asn Tyr Trp Arg Lys His Asn Ser Ala Ile Ser Ala Val Ile Tyr Asn
    130                 135                 140

gat gat ggt ctt gat gtt gcc aat gaa aag atc aga caa cta ttt att    480
Asp Asp Gly Leu Asp Val Ala Asn Glu Lys Ile Arg Gln Leu Phe Ile
145                 150                 155                 160

ggt cgt tat tta tca ttt acc gga ggt aac aca tta act cag atg gaa    528
Gly Arg Tyr Leu Ser Phe Thr Gly Gly Asn Thr Leu Thr Gln Met Glu
                165                 170                 175

ttc acc ata atg ggg tat atg gtt tct ggt tat aat cca tat cag att    576
Phe Thr Ile Met Gly Tyr Met Val Ser Gly Tyr Asn Pro Tyr Gln Ile
            180                 185                 190

gct gaa gtt ctg gat atg gat atc cgt agc atc tac gcg tac aag caa    624
Ala Glu Val Leu Asp Met Asp Ile Arg Ser Ile Tyr Ala Tyr Lys Gln
        195                 200                 205

cga atc gaa aag aga atg ggt ggt aaa ata aac gaa tta ttt att cgt    672
Arg Ile Glu Lys Arg Met Gly Gly Lys Ile Asn Glu Leu Phe Ile Arg
    210                 215                 220

tca cat tcg gtc caa cat tga                                        693
Ser His Ser Val Gln His  *
225                 230


<210> 180
<211> 1176
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1176)

<400> 180
atg caa aaa ctc att aac tca gtg caa aac tat gcc tgg ggc agc aaa    48
Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
 1               5                  10                  15

acg gcg ttg act gaa ctt tat ggt atg gaa aat ccg tcc agc cag ccg    96
Thr Ala Leu Thr Glu Leu Tyr Gly Met Glu Asn Pro Ser Ser Gln Pro
            20                  25                  30

atg gcc gag ctg tgg atg ggc gca cat ccg aaa agc agt tca cga gtg    144
Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Arg Val
        35                  40                  45
```

```
cag aat gcc gcc gga gat atc gtt tca ctg cgt gat gtg att gag agt    192
Gln Asn Ala Ala Gly Asp Ile Val Ser Leu Arg Asp Val Ile Glu Ser
     50                  55                  60

gat aaa tcg act ctg ctc gga gag gcc gtt gcc aaa cgc ttt ggc gaa    240
Asp Lys Ser Thr Leu Leu Gly Glu Ala Val Ala Lys Arg Phe Gly Glu
 65                  70                  75                  80

ctg cct ttc ctg ttc aaa gta tta tgc gca gca cag cca ctc tcc att    288
Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Ala Gln Pro Leu Ser Ile
                 85                  90                  95

cag gtt cat cca aac aaa cac aat tct gaa atc ggt ttt gcc aaa gaa    336
Gln Val His Pro Asn Lys His Asn Ser Glu Ile Gly Phe Ala Lys Glu
            100                 105                 110

aat gcc gca ggt atc ccg atg gat gcc gcc gag cgt aac tat aaa gat    384
Asn Ala Ala Gly Ile Pro Met Asp Ala Ala Glu Arg Asn Tyr Lys Asp
        115                 120                 125

cct aac cac aag ccg gag ctg gtt ttt gcg ctg acg cct ttc ctt gcg    432
Pro Asn His Lys Pro Glu Leu Val Phe Ala Leu Thr Pro Phe Leu Ala
    130                 135                 140

atg aac gcg ttt cgt gaa ttt tcc gag att gtc tcc cta ctc cag ccg    480
Met Asn Ala Phe Arg Glu Phe Ser Glu Ile Val Ser Leu Leu Gln Pro
145                 150                 155                 160

gtc gca ggt gca cat ccg gcg att gct cac ttt tta caa cag cct gat    528
Val Ala Gly Ala His Pro Ala Ile Ala His Phe Leu Gln Gln Pro Asp
                165                 170                 175

gcc gaa cgt tta agc gaa ctg ttc gcc agc ctg ttg aat atg cag ggt    576
Ala Glu Arg Leu Ser Glu Leu Phe Ala Ser Leu Leu Asn Met Gln Gly
            180                 185                 190

gaa gaa aaa tcc cgc gcg ctg gcg att tta aaa tcg gcc ctc gat agc    624
Glu Glu Lys Ser Arg Ala Leu Ala Ile Leu Lys Ser Ala Leu Asp Ser
        195                 200                 205

cag cag ggt gaa ccg tgg caa acg att cgt tta att tct gaa ttt tac    672
Gln Gln Gly Glu Pro Trp Gln Thr Ile Arg Leu Ile Ser Glu Phe Tyr
    210                 215                 220

ccg gaa gac agc ggt ctg ttc tcc ccg cta ttg ctg aat gtg gtg aaa    720
Pro Glu Asp Ser Gly Leu Phe Ser Pro Leu Leu Leu Asn Val Val Lys
225                 230                 235                 240

ttg aac cct ggc gaa gcg atg ttc ctg ttc gct gaa aca ccg cac gct    768
Leu Asn Pro Gly Glu Ala Met Phe Leu Phe Ala Glu Thr Pro His Ala
                245                 250                 255

tac ctg caa ggc gtg gcg ctg gaa gtg atg gca aac tcc gat aac gtg    816
Tyr Leu Gln Gly Val Ala Leu Glu Val Met Ala Asn Ser Asp Asn Val
            260                 265                 270

ctg cgt gcg ggt ctg acg cct aaa tac att gat att ccg gaa ctg gtt    864
Leu Arg Ala Gly Leu Thr Pro Lys Tyr Ile Asp Ile Pro Glu Leu Val
        275                 280                 285
```

gcc aat gtg aaa ttc gaa gcc aaa ccg gct aac cag ttg ttg acc cag 912
Ala Asn Val Lys Phe Glu Ala Lys Pro Ala Asn Gln Leu Leu Thr Gln
290 295 300

ccg gtg aaa caa ggt gca gaa ctg gac ttc ccg att cca gtg gat gat 960
Pro Val Lys Gln Gly Ala Glu Leu Asp Phe Pro Ile Pro Val Asp Asp
305 310 315 320

ttt gcc ttc tcg ctg cat gac ctt agt gat aaa gaa acc acc att agc 1008
Phe Ala Phe Ser Leu His Asp Leu Ser Asp Lys Glu Thr Thr Ile Ser
325 330 335

cag cag agt gcc gcc att ttg ttc tgc gtc gaa ggc gat gca acg ttg 1056
Gln Gln Ser Ala Ala Ile Leu Phe Cys Val Glu Gly Asp Ala Thr Leu
340 345 350

tgg aaa ggt tct cag cag tta cag ctt aaa ccg ggt gaa tca gcg ttt 1104
Trp Lys Gly Ser Gln Gln Leu Gln Leu Lys Pro Gly Glu Ser Ala Phe
355 360 365

att gcc gcc aac gaa tca ccg gtg act gtc aaa ggc cac ggc cgt tta 1152
Ile Ala Ala Asn Glu Ser Pro Val Thr Val Lys Gly His Gly Arg Leu
370 375 380

gcg cgt gtt tac aac aag ctg taa 1176
Ala Arg Val Tyr Asn Lys Leu *
385 390

<210> 181
<211> 795
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(795)

<400> 181
atg aaa cag tat tta gaa ctg atg caa aaa gtg ctc gac gaa ggc aca 48
Met Lys Gln Tyr Leu Glu Leu Met Gln Lys Val Leu Asp Glu Gly Thr
1 5 10 15

cag aaa aac gac cgt acc gga acc gga acg ctt tcc att ttt ggt cat 96
Gln Lys Asn Asp Arg Thr Gly Thr Gly Thr Leu Ser Ile Phe Gly His
20 25 30

cag atg cgt ttt aac ctg caa gat gga ttc ccg ctg gtg aca act aaa 144
Gln Met Arg Phe Asn Leu Gln Asp Gly Phe Pro Leu Val Thr Thr Lys
35 40 45

cgt tgc cac ctg cgt tcc atc atc cat gaa ctg ctg tgg ttt ctg cag 192
Arg Cys His Leu Arg Ser Ile Ile His Glu Leu Leu Trp Phe Leu Gln
50 55 60

ggc gac act aac att gct tat cta cac gaa aac aat gtc acc atc tgg 240
Gly Asp Thr Asn Ile Ala Tyr Leu His Glu Asn Asn Val Thr Ile Trp
65 70 75 80

gac gaa tgg gcc gat gaa aac ggc gac ctc ggg cca gtg tat ggt aaa 288
Asp Glu Trp Ala Asp Glu Asn Gly Asp Leu Gly Pro Val Tyr Gly Lys

```
                85                  90                  95

cag tgg cgc gcc tgg cca acg cca gat ggt cgt cat att gac cag atc    336
Gln Trp Arg Ala Trp Pro Thr Pro Asp Gly Arg His Ile Asp Gln Ile
            100                 105                 110

act acg gta ctg aac cag ctg aaa aac gac ccg gat tcg cgc cgc att    384
Thr Thr Val Leu Asn Gln Leu Lys Asn Asp Pro Asp Ser Arg Arg Ile
        115                 120                 125

att gtt tca gcg tgg aac gta ggc gaa ctg gat aaa atg gcg ctg gca    432
Ile Val Ser Ala Trp Asn Val Gly Glu Leu Asp Lys Met Ala Leu Ala
    130                 135                 140

ccg tgc cat gca ttc ttc cag ttc tat gtg gca gac ggc aaa ctc tct    480
Pro Cys His Ala Phe Phe Gln Phe Tyr Val Ala Asp Gly Lys Leu Ser
145                 150                 155                 160

tgc cag ctt tat cag cgc tcc tgt gac gtc ttc ctc ggc ctg ccg ttc    528
Cys Gln Leu Tyr Gln Arg Ser Cys Asp Val Phe Leu Gly Leu Pro Phe
                165                 170                 175

aac att gcc agc tac gcg tta ttg gtg cat atg atg gcg cag cag tgc    576
Asn Ile Ala Ser Tyr Ala Leu Leu Val His Met Met Ala Gln Gln Cys
            180                 185                 190

gat ctg gaa gtg ggt gat ttt gtc tgg acc ggt ggc gac acg cat ctg    624
Asp Leu Glu Val Gly Asp Phe Val Trp Thr Gly Gly Asp Thr His Leu
        195                 200                 205

tac agc aac cat atg gat caa act cat ctg caa tta agc cgc gaa ccg    672
Tyr Ser Asn His Met Asp Gln Thr His Leu Gln Leu Ser Arg Glu Pro
    210                 215                 220

cgt ccg ctg ccg aag ttg att atc aaa cgt aaa ccc gaa tcc atc ttc    720
Arg Pro Leu Pro Lys Leu Ile Ile Lys Arg Lys Pro Glu Ser Ile Phe
225                 230                 235                 240

gac tac cgt ttc gaa gac ttt gag att gaa ggc tac gat ccg cat ccg    768
Asp Tyr Arg Phe Glu Asp Phe Glu Ile Glu Gly Tyr Asp Pro His Pro
                245                 250                 255

ggc att aaa gcg ccg gtg gct atc taa                                795
Gly Ile Lys Ala Pro Val Ala Ile  *
            260
```

<210> 182
<211> 876
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(876)

<400> 182

```
atg acc agt agc tat ctg cat ttt ccg gag ttt gat ccg gtc att ttc    48
Met Thr Ser Ser Tyr Leu His Phe Pro Glu Phe Asp Pro Val Ile Phe
 1               5                  10                  15
```

tca ata gga ccc gtg gcg ctt cac tgg tac ggc ctg atg tat ctg gtg 96
Ser Ile Gly Pro Val Ala Leu His Trp Tyr Gly Leu Met Tyr Leu Val
20 25 30

ggt ttc att ttt gca atg tgg ctg gca aca cga cgg gcg aat cgt ccg 144
Gly Phe Ile Phe Ala Met Trp Leu Ala Thr Arg Arg Ala Asn Arg Pro
35 40 45

ggc agc ggc tgg acc aaa aat gaa gtt gaa aac tta ctc tat gcg ggc 192
Gly Ser Gly Trp Thr Lys Asn Glu Val Glu Asn Leu Leu Tyr Ala Gly
50 55 60

ttc ctc ggc gtc ttc ctc ggg gga cgt att ggt tat gtt ctg ttc tac 240
Phe Leu Gly Val Phe Leu Gly Gly Arg Ile Gly Tyr Val Leu Phe Tyr
65 70 75 80

aat ttc ccg cag ttt atg gcc gat ccg ctg tat ctg ttc cgt gtc tgg 288
Asn Phe Pro Gln Phe Met Ala Asp Pro Leu Tyr Leu Phe Arg Val Trp
85 90 95

gac ggc ggc atg tct ttc cac ggc ggc ctg att ggc gtt atc gtg gtg 336
Asp Gly Gly Met Ser Phe His Gly Gly Leu Ile Gly Val Ile Val Val
100 105 110

atg att atc ttc gcc cgc cgt act aaa cgt tcc ttc ttc cag gtc tct 384
Met Ile Ile Phe Ala Arg Arg Thr Lys Arg Ser Phe Phe Gln Val Ser
115 120 125

gat ttt atc gca cca ctc att ccg ttt ggt ctt ggt gcc ggg cgt ctg 432
Asp Phe Ile Ala Pro Leu Ile Pro Phe Gly Leu Gly Ala Gly Arg Leu
130 135 140

ggc aac ttt att aac ggt gaa ttg tgg ggc cgc gtt gac ccg aac ttc 480
Gly Asn Phe Ile Asn Gly Glu Leu Trp Gly Arg Val Asp Pro Asn Phe
145 150 155 160

ccg ttt gcc atg ctg ttc cct ggc tcc cgt aca gaa gat att ttg ctg 528
Pro Phe Ala Met Leu Phe Pro Gly Ser Arg Thr Glu Asp Ile Leu Leu
165 170 175

ctg caa acc aac ccg cag tgg caa tcc att ttc gac act tac ggt gtg 576
Leu Gln Thr Asn Pro Gln Trp Gln Ser Ile Phe Asp Thr Tyr Gly Val
180 185 190

ctg ccg cgc cac cca tca cag ctt tac gag ctg ctg ctg gaa ggt gtg 624
Leu Pro Arg His Pro Ser Gln Leu Tyr Glu Leu Leu Leu Glu Gly Val
195 200 205

gtg ctg ttt att atc ctc aac ctg tat att cgt aaa cca cgc cca atg 672
Val Leu Phe Ile Ile Leu Asn Leu Tyr Ile Arg Lys Pro Arg Pro Met
210 215 220

gga gct gtc tca ggt ttg ttc ctg att ggt tac ggc gcg ttt cgc atc 720
Gly Ala Val Ser Gly Leu Phe Leu Ile Gly Tyr Gly Ala Phe Arg Ile
225 230 235 240

att gtt gag ttt ttc cgc cag ccc gac gcg cag ttt acc ggt gcc tgg 768
Ile Val Glu Phe Phe Arg Gln Pro Asp Ala Gln Phe Thr Gly Ala Trp
245 250 255

gtg cag tac atc agc atg ggg caa att ctt tcc atc ccg atg att gtc 816

Val Gln Tyr Ile Ser Met Gly Gln Ile Leu Ser Ile Pro Met Ile Val
260 265 270

gcg ggt gtg atc atg atg gtc tgg gca tat cgt cgc agc cca cag caa 864
Ala Gly Val Ile Met Met Val Trp Ala Tyr Arg Arg Ser Pro Gln Gln
275 280 285

cac gtt tcc tga 876
His Val Ser *
290

<210> 183
<211> 726
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(726)

<400> 183
atg gac agt ctc aat ctt aat aaa cat att tcc ggc cag ttc aac gcc 48
Met Asp Ser Leu Asn Leu Asn Lys His Ile Ser Gly Gln Phe Asn Ala
1 5 10 15

gaa ctg gaa agt atc cgc acg cag gtg atg acc atg ggc ggc atg gtg 96
Glu Leu Glu Ser Ile Arg Thr Gln Val Met Thr Met Gly Gly Met Val
20 25 30

gag cag cag ctt tct gat gca atc acc gcg atg cat aac cag gac agc 144
Glu Gln Gln Leu Ser Asp Ala Ile Thr Ala Met His Asn Gln Asp Ser
35 40 45

gat ctg gcg aag cgc gtc atc gaa ggc gac aag aac gtc aac atg atg 192
Asp Leu Ala Lys Arg Val Ile Glu Gly Asp Lys Asn Val Asn Met Met
50 55 60

gaa gtg gcg atc gat gaa gcc tgc gtg cgc att atc gcc aaa cgt cag 240
Glu Val Ala Ile Asp Glu Ala Cys Val Arg Ile Ile Ala Lys Arg Gln
65 70 75 80

ccg acg gcg agc gac ctg cga ctg gtt atg gtg atc agt aaa acc att 288
Pro Thr Ala Ser Asp Leu Arg Leu Val Met Val Ile Ser Lys Thr Ile
85 90 95

gcc gag ctg gag cgt att ggc gac gtg gcg gac aaa atc tgc cgt act 336
Ala Glu Leu Glu Arg Ile Gly Asp Val Ala Asp Lys Ile Cys Arg Thr
100 105 110

gcg ctg gag aaa ttc tcc cag cag cat cag ccg ttg ctg gta agt ctg 384
Ala Leu Glu Lys Phe Ser Gln Gln His Gln Pro Leu Leu Val Ser Leu
115 120 125

gag tcg ctg ggc cgt cat acc atc cag atg ctg cac gac gtg ctg gac 432
Glu Ser Leu Gly Arg His Thr Ile Gln Met Leu His Asp Val Leu Asp
130 135 140

gcg ttc gcg cgg atg gac att gac gaa gcg gta cgt att tat cgt gaa 480
Ala Phe Ala Arg Met Asp Ile Asp Glu Ala Val Arg Ile Tyr Arg Glu
145 150 155 160

```
gat aaa aaa gtc gat cag gaa tac gaa ggt att gtt cgt caa ctg atg    528
Asp Lys Lys Val Asp Gln Glu Tyr Glu Gly Ile Val Arg Gln Leu Met
                165                 170                 175

acc tac atg atg gaa gat tcg cgt acc att ccg agc gta ctt act gcg    576
Thr Tyr Met Met Glu Asp Ser Arg Thr Ile Pro Ser Val Leu Thr Ala
            180                 185                 190

ctg ttc tgc gcg cgt tct atc gaa cgt att ggc gac cgc tgc cag aat    624
Leu Phe Cys Ala Arg Ser Ile Glu Arg Ile Gly Asp Arg Cys Gln Asn
        195                 200                 205

att tgt gag ttt atc ttc tac tac gtg aag ggg cag gat ttc cgt cac    672
Ile Cys Glu Phe Ile Phe Tyr Tyr Val Lys Gly Gln Asp Phe Arg His
    210                 215                 220

gtc ggt ggc gat gag ctg gat aaa ctg ctg gcg ggg aaa gat agc gac    720
Val Gly Gly Asp Glu Leu Asp Lys Leu Leu Ala Gly Lys Asp Ser Asp
225                 230                 235                 240

aaa taa                                                            726
Lys  *
```

<210> 184
<211> 774
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(774)

<400> 184

```
atg agt atg gtt gaa act gcc ccg agt aaa att cag gtt cgt aat ttg    48
Met Ser Met Val Glu Thr Ala Pro Ser Lys Ile Gln Val Arg Asn Leu
 1               5                  10                  15

aac ttc tac tac ggc aaa ttc cat gcc ctg aaa aac atc aac ctg gat    96
Asn Phe Tyr Tyr Gly Lys Phe His Ala Leu Lys Asn Ile Asn Leu Asp
            20                  25                  30

atc gct aaa aac cag gta acg gcg ttt atc ggg ccg tcc ggc tgc ggt    144
Ile Ala Lys Asn Gln Val Thr Ala Phe Ile Gly Pro Ser Gly Cys Gly
        35                  40                  45

aaa tcg acg ctg ctg cgt acc ttc aac aaa atg ttt gaa ctg tac ccg    192
Lys Ser Thr Leu Leu Arg Thr Phe Asn Lys Met Phe Glu Leu Tyr Pro
    50                  55                  60

gag cag cgt gcg gaa ggt gaa att ctg ctt gat ggc gac aac atc ctg    240
Glu Gln Arg Ala Glu Gly Glu Ile Leu Leu Asp Gly Asp Asn Ile Leu
65                  70                  75                  80

acc aac tct cag gat atc gca ctg ctg cgt gcg aaa gtg ggc atg gtg    288
Thr Asn Ser Gln Asp Ile Ala Leu Leu Arg Ala Lys Val Gly Met Val
                85                  90                  95

ttc cag aaa ccg acg ccg ttt ccg atg tcc atc tac gac aac atc gct    336
```

```
Phe Gln Lys Pro Thr Pro Phe Pro Met Ser Ile Tyr Asp Asn Ile Ala
            100                 105                 110

ttt ggc gtt cgt ctg ttt gag aag ctc tcc cgt gcc gac atg gac gag      384
Phe Gly Val Arg Leu Phe Glu Lys Leu Ser Arg Ala Asp Met Asp Glu
        115                 120                 125

cgc gtg cag tgg gca ttg acc aaa gcc gca ttg tgg aac gaa acc aaa      432
Arg Val Gln Trp Ala Leu Thr Lys Ala Ala Leu Trp Asn Glu Thr Lys
    130                 135                 140

gat aaa ttg cac cag agc ggt tac tct ctc tct ggt ggt cag caa cag      480
Asp Lys Leu His Gln Ser Gly Tyr Ser Leu Ser Gly Gly Gln Gln Gln
145                 150                 155                 160

cgt ctg tgt att gcg cgt ggt atc gcc att cgc ccg gaa gtg ctg ctg      528
Arg Leu Cys Ile Ala Arg Gly Ile Ala Ile Arg Pro Glu Val Leu Leu
                165                 170                 175

ctc gac gaa ccg tgt tcg gcg ctc gac cct atc tct acc ggg cgt att      576
Leu Asp Glu Pro Cys Ser Ala Leu Asp Pro Ile Ser Thr Gly Arg Ile
            180                 185                 190

gaa gag ctg atc acc gaa ctg aag cag gat tac acc gtg gtg atc gtc      624
Glu Glu Leu Ile Thr Glu Leu Lys Gln Asp Tyr Thr Val Val Ile Val
        195                 200                 205

acc cac aac atg cag cag gct gcg cgt tgt tcc gac cac acg gcg ttt      672
Thr His Asn Met Gln Gln Ala Ala Arg Cys Ser Asp His Thr Ala Phe
    210                 215                 220

atg tac ctg ggc gaa ttg att gag ttc agc aac acg gac gat ctg ttc      720
Met Tyr Leu Gly Glu Leu Ile Glu Phe Ser Asn Thr Asp Asp Leu Phe
225                 230                 235                 240

acc aag cca gcg aag aaa caa aca gaa gac tac atc acc ggt cgt tac      768
Thr Lys Pro Ala Lys Lys Gln Thr Glu Asp Tyr Ile Thr Gly Arg Tyr
                245                 250                 255

ggt tga                                                              774
Gly  *
```

<210> 185
<211> 891
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(891)

<400> 185

```
atg gct atg gtt gaa atg caa acc act gcg gcg ctg gct gaa tct cgc      48
Met Ala Met Val Glu Met Gln Thr Thr Ala Ala Leu Ala Glu Ser Arg
 1               5                  10                  15

cgc aaa atg cag gcg cgt cgc cgc ctc aaa aac cgt att gcg ctg acg      96
Arg Lys Met Gln Ala Arg Arg Arg Leu Lys Asn Arg Ile Ala Leu Thr
            20                  25                  30
```

```
ctc tcg atg gcg acg atg gcc ttc ggc ctg ttc tgg ctg atc tgg att    144
Leu Ser Met Ala Thr Met Ala Phe Gly Leu Phe Trp Leu Ile Trp Ile
        35                  40                  45

tta atg tcc acc atc act cgc ggt atc gac ggt atg tcg ctg gcg ctg    192
Leu Met Ser Thr Ile Thr Arg Gly Ile Asp Gly Met Ser Leu Ala Leu
    50                  55                  60

ttc act gaa atg acg ccg ccg ccc aat acg gaa ggt ggt ggt ctg gcg    240
Phe Thr Glu Met Thr Pro Pro Pro Asn Thr Glu Gly Gly Gly Leu Ala
65                  70                  75                  80

aac gct ctg gcg ggt agc ggg ctg tta att ttg tgg gcc acg gta ttc    288
Asn Ala Leu Ala Gly Ser Gly Leu Leu Ile Leu Trp Ala Thr Val Phe
                85                  90                  95

ggt acg ccg ctg ggc att atg gcg ggg att tat ctg gcg gaa tat ggt    336
Gly Thr Pro Leu Gly Ile Met Ala Gly Ile Tyr Leu Ala Glu Tyr Gly
            100                 105                 110

cgt aaa tcc tgg ctg gca gaa gtg att cgc ttc att aac gac att ctg    384
Arg Lys Ser Trp Leu Ala Glu Val Ile Arg Phe Ile Asn Asp Ile Leu
        115                 120                 125

ctc tct gcg ccg tcg att gtg gtt ggt ctg ttt gtt tac acc att gtg    432
Leu Ser Ala Pro Ser Ile Val Val Gly Leu Phe Val Tyr Thr Ile Val
    130                 135                 140

gtg gcg cag atg gag cac ttc tcc ggc tgg gcg ggc gtg att gcc ctg    480
Val Ala Gln Met Glu His Phe Ser Gly Trp Ala Gly Val Ile Ala Leu
145                 150                 155                 160

gcg ttg ttg cag gtg ccg att gtt atc cgc acc acc gag aac atg ctg    528
Ala Leu Leu Gln Val Pro Ile Val Ile Arg Thr Thr Glu Asn Met Leu
                165                 170                 175

aaa ctg gtg ccg tac agc ctg cgt gaa gcg gct tat gcg ctg ggt aca    576
Lys Leu Val Pro Tyr Ser Leu Arg Glu Ala Ala Tyr Ala Leu Gly Thr
            180                 185                 190

ccg aag tgg aag atg atc tct gcg att acg ctg aaa gcg tcg gtg tcc    624
Pro Lys Trp Lys Met Ile Ser Ala Ile Thr Leu Lys Ala Ser Val Ser
        195                 200                 205

ggg att atg acc ggt atc ctg ctg gcg att gcc cgt att gct ggt gaa    672
Gly Ile Met Thr Gly Ile Leu Leu Ala Ile Ala Arg Ile Ala Gly Glu
    210                 215                 220

acc gcg ccg ctg ctg ttt acc gcg ctc tcc aac cag ttc tgg agc acg    720
Thr Ala Pro Leu Leu Phe Thr Ala Leu Ser Asn Gln Phe Trp Ser Thr
225                 230                 235                 240

gac atg atg cag ccg atc gcc aac ctg ccg gtg acg atc ttt aag ttt    768
Asp Met Met Gln Pro Ile Ala Asn Leu Pro Val Thr Ile Phe Lys Phe
                245                 250                 255

gcg atg agc ccg ttt gcg gaa tgg cag caa ttg gcc tgg gcc ggg gta    816
Ala Met Ser Pro Phe Ala Glu Trp Gln Gln Leu Ala Trp Ala Gly Val
            260                 265                 270
```

ttg atc att acc ctg tgc gta ctg ctg ctg aac att ctg gcg cgc gtt 864
Leu Ile Ile Thr Leu Cys Val Leu Leu Leu Asn Ile Leu Ala Arg Val
275 280 285

gtt ttt gcg aag aat aaa cac ggt tga 891
Val Phe Ala Lys Asn Lys His Gly *
290 295

<210> 186
<211> 960
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(960)

<400> 186
atg gct gca acc aag cct gct ttt aac cca ccg ggt aaa aag ggc gac 48
Met Ala Ala Thr Lys Pro Ala Phe Asn Pro Pro Gly Lys Lys Gly Asp
1 5 10 15

ata att ttc agc gtg ctg gta aaa ctg gcg gcg ctg att gtg cta ttg 96
Ile Ile Phe Ser Val Leu Val Lys Leu Ala Ala Leu Ile Val Leu Leu
20 25 30

atg ttg ggt ggc att att gtc tct ctg atc atc tcc tcc tgg ccg agc 144
Met Leu Gly Gly Ile Ile Val Ser Leu Ile Ile Ser Ser Trp Pro Ser
35 40 45

att cag aaa ttt ggt ctg gct ttc cta tgg acc aaa gag tgg gat gca 192
Ile Gln Lys Phe Gly Leu Ala Phe Leu Trp Thr Lys Glu Trp Asp Ala
50 55 60

ccg aac gat atc tac ggg gcg ctg gtg ccg atc tac ggt acg ttg gtg 240
Pro Asn Asp Ile Tyr Gly Ala Leu Val Pro Ile Tyr Gly Thr Leu Val
65 70 75 80

act tcg ttt atc gcg ctg ctg atc gcc gtc ccg gtg agt ttc ggt atc 288
Thr Ser Phe Ile Ala Leu Leu Ile Ala Val Pro Val Ser Phe Gly Ile
85 90 95

gcc ctg ttc ctg act gag ctt gcg cct ggc tgg ctg aaa cgc ccg ctg 336
Ala Leu Phe Leu Thr Glu Leu Ala Pro Gly Trp Leu Lys Arg Pro Leu
100 105 110

ggt atc gcc att gag ctg ctg gca gcc att cca agt atc gtt tac ggc 384
Gly Ile Ala Ile Glu Leu Leu Ala Ala Ile Pro Ser Ile Val Tyr Gly
115 120 125

atg tgg ggc ctg ttt atc ttt gcg ccg ctg ttc gcc gtt tac ttt cag 432
Met Trp Gly Leu Phe Ile Phe Ala Pro Leu Phe Ala Val Tyr Phe Gln
130 135 140

gag ccg gtc ggc aat atc atg tcg aat atc ccg att gtt ggc gcg ctg 480
Glu Pro Val Gly Asn Ile Met Ser Asn Ile Pro Ile Val Gly Ala Leu
145 150 155 160

ttc tct ggc ccc gca ttt ggt atc ggt atc ctc gcg gca ggc gtg atc 528
Phe Ser Gly Pro Ala Phe Gly Ile Gly Ile Leu Ala Ala Gly Val Ile

165 170 175

ctc gcc atc atg att att ccg tac att gcg gcg gta atg cgt gat gtg 576
Leu Ala Ile Met Ile Ile Pro Tyr Ile Ala Ala Val Met Arg Asp Val
180 185 190

ttc gaa caa acc ccg gtg atg atg aaa gag tcg gcc tac ggt att ggc 624
Phe Glu Gln Thr Pro Val Met Met Lys Glu Ser Ala Tyr Gly Ile Gly
195 200 205

tgc acc acc tgg gaa gtt atc tgg cgt atc gtt ctt ccg ttc acc aaa 672
Cys Thr Thr Trp Glu Val Ile Trp Arg Ile Val Leu Pro Phe Thr Lys
210 215 220

aat ggt gtt atc ggc ggc atc atg ctg ggg ctg ggc cgc gcg ctc ggt 720
Asn Gly Val Ile Gly Gly Ile Met Leu Gly Leu Gly Arg Ala Leu Gly
225 230 235 240

gaa acc atg gcg gtg acc ttt atc atc ggt aac acc tac cag ctc gac 768
Glu Thr Met Ala Val Thr Phe Ile Ile Gly Asn Thr Tyr Gln Leu Asp
245 250 255

agc gcc tcg ctg tat atg ccg ggc aac agt atc acc tct gcg ctg gcg 816
Ser Ala Ser Leu Tyr Met Pro Gly Asn Ser Ile Thr Ser Ala Leu Ala
260 265 270

aac gaa ttt gcg gaa gcg gaa tcc ggt ctg cac gtt gcc gca ctg atg 864
Asn Glu Phe Ala Glu Ala Glu Ser Gly Leu His Val Ala Ala Leu Met
275 280 285

gaa ctg ggc ctg atc ctg ttt gtg att acc ttc atc gtc ctc gcc gca 912
Glu Leu Gly Leu Ile Leu Phe Val Ile Thr Phe Ile Val Leu Ala Ala
290 295 300

tcg aag ttt atg att atg cgc ctg gct aag aat gag ggg gca cgc taa 960
Ser Lys Phe Met Ile Met Arg Leu Ala Lys Asn Glu Gly Ala Arg *
305 310 315

<210> 187
<211> 1041
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1041)

<400> 187
atg aaa gtt atg cgt acc acc gtc gca act gtt gtc gcc gcg acc tta 48
Met Lys Val Met Arg Thr Thr Val Ala Thr Val Val Ala Ala Thr Leu
1 5 10 15

tcg atg agt gct ttc tct gtg ttt gca gaa gca agc ctg aca ggt gca 96
Ser Met Ser Ala Phe Ser Val Phe Ala Glu Ala Ser Leu Thr Gly Ala
20 25 30

ggt gca acc ttc cct gcg ccg gtg tat gcc aaa tgg gct gac act tac 144
Gly Ala Thr Phe Pro Ala Pro Val Tyr Ala Lys Trp Ala Asp Thr Tyr
35 40 45

```
cag aaa gaa acc ggt aat aaa gtt aac tac cag ggt atc ggt tct tcc     192
Gln Lys Glu Thr Gly Asn Lys Val Asn Tyr Gln Gly Ile Gly Ser Ser
     50                  55                  60

ggt ggc gta aaa cag att atc gct aat acc gtt gat ttt ggt gcc tct     240
Gly Gly Val Lys Gln Ile Ile Ala Asn Thr Val Asp Phe Gly Ala Ser
 65                  70                  75                  80

gac gcg ccg ctg tct gac gaa aaa ctg gct cag gaa ggt ctg ttc cag     288
Asp Ala Pro Leu Ser Asp Glu Lys Leu Ala Gln Glu Gly Leu Phe Gln
                 85                  90                  95

ttc ccg acc gtg att ggc ggc gtg gtg ctg gcg gtt aac att cca ggg     336
Phe Pro Thr Val Ile Gly Gly Val Val Leu Ala Val Asn Ile Pro Gly
            100                 105                 110

ctg aag tct ggc gaa ctg gtg ctg gat ggt aaa acc ctc ggc gac atc     384
Leu Lys Ser Gly Glu Leu Val Leu Asp Gly Lys Thr Leu Gly Asp Ile
        115                 120                 125

tac ctg ggc aaa atc aag aag tgg gat gat gaa gcc atc gcc aaa ctg     432
Tyr Leu Gly Lys Ile Lys Lys Trp Asp Asp Glu Ala Ile Ala Lys Leu
    130                 135                 140

aat ccg ggt ctg aaa ctg cct tca caa aac att gct gta gta cgc cgc     480
Asn Pro Gly Leu Lys Leu Pro Ser Gln Asn Ile Ala Val Val Arg Arg
145                 150                 155                 160

gca gat ggc tcc ggg act tcc ttc gtc ttc acc agc tac ctg gcg aaa     528
Ala Asp Gly Ser Gly Thr Ser Phe Val Phe Thr Ser Tyr Leu Ala Lys
                165                 170                 175

gtg aac gaa gag tgg aaa aac aac gtt ggt act ggc tct acc gta aaa     576
Val Asn Glu Glu Trp Lys Asn Asn Val Gly Thr Gly Ser Thr Val Lys
            180                 185                 190

tgg ccg atc ggt ctg ggc ggt aaa ggt aac gac ggt atc gcc gcg ttc     624
Trp Pro Ile Gly Leu Gly Gly Lys Gly Asn Asp Gly Ile Ala Ala Phe
        195                 200                 205

gtt cag cgt ctg ccg ggt gca att ggt tat gtt gaa tat gct tac gcg     672
Val Gln Arg Leu Pro Gly Ala Ile Gly Tyr Val Glu Tyr Ala Tyr Ala
    210                 215                 220

aag cag aac aac ctg gcg tac acc aaa ctg atc tcc gct gat ggt aaa     720
Lys Gln Asn Asn Leu Ala Tyr Thr Lys Leu Ile Ser Ala Asp Gly Lys
225                 230                 235                 240

ccg gtt agt ccg acc gaa gaa aac ttc gct aat gca gca aaa ggt gca     768
Pro Val Ser Pro Thr Glu Glu Asn Phe Ala Asn Ala Ala Lys Gly Ala
                245                 250                 255

gac tgg agc aaa acc ttc gct cag gat ctg acc aac cag aaa ggc gaa     816
Asp Trp Ser Lys Thr Phe Ala Gln Asp Leu Thr Asn Gln Lys Gly Glu
            260                 265                 270

gat gca tgg cct att acc tct acc acg ttc att ctg atc cac aaa gat     864
Asp Ala Trp Pro Ile Thr Ser Thr Thr Phe Ile Leu Ile His Lys Asp
        275                 280                 285
```

```
cag aag aaa cca gaa caa ggc aca gaa gtg ctg aaa ttc ttc gac tgg    912
Gln Lys Lys Pro Glu Gln Gly Thr Glu Val Leu Lys Phe Phe Asp Trp
    290                 295                 300

gcg tac aaa acc ggg gct aaa cag gcg aac gac ctg gat tac gcc agc    960
Ala Tyr Lys Thr Gly Ala Lys Gln Ala Asn Asp Leu Asp Tyr Ala Ser
305                 310                 315                 320

ctg ccg gat agt gta gtt gaa cag gtt cgc gct gcg tgg aag acc aat    1008
Leu Pro Asp Ser Val Val Glu Gln Val Arg Ala Ala Trp Lys Thr Asn
                325                 330                 335

att aaa gac agt agc ggt aag ccg ctg tac taa                        1041
Ile Lys Asp Ser Ser Gly Lys Pro Leu Tyr  *
            340                 345
```

<210> 188
<211> 711
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(711)

<400> 188

```
atg gga tct gga ctg gtg aac gga ggt gat tat ttt tat aat aac ctt    48
Met Gly Ser Gly Leu Val Asn Gly Gly Asp Tyr Phe Tyr Asn Asn Leu
 1               5                  10                  15

tca ttc aca gtc acc agg tac aac ggc atc atg gca act gac tca aca    96
Ser Phe Thr Val Thr Arg Tyr Asn Gly Ile Met Ala Thr Asp Ser Thr
            20                  25                  30

caa tgt gta aaa aaa agc cgt ggc cgc cca aaa gtg ttc gac agg gat    144
Gln Cys Val Lys Lys Ser Arg Gly Arg Pro Lys Val Phe Asp Arg Asp
        35                  40                  45

gcc gcg ctt gat aag gcc atg aaa ttg ttc tgg caa cac ggt tat gaa    192
Ala Ala Leu Asp Lys Ala Met Lys Leu Phe Trp Gln His Gly Tyr Glu
    50                  55                  60

gcg act tct ctt gcg gac ctc gtc gaa gcg acc gga gcc aaa gcg ccc    240
Ala Thr Ser Leu Ala Asp Leu Val Glu Ala Thr Gly Ala Lys Ala Pro
65                  70                  75                  80

acg cta tac gcg gaa ttt acc aac aaa gag ggg tta ttt cgt gcc gtt    288
Thr Leu Tyr Ala Glu Phe Thr Asn Lys Glu Gly Leu Phe Arg Ala Val
                85                  90                  95

ctc gac cgc tat atc gat cgt ttt gcc gct aag cat gaa gca cag ctg    336
Leu Asp Arg Tyr Ile Asp Arg Phe Ala Ala Lys His Glu Ala Gln Leu
            100                 105                 110

ttt tgt gaa gag aaa agc gtg gag tct gcg ctg gct gac tat ttt gct    384
Phe Cys Glu Glu Lys Ser Val Glu Ser Ala Leu Ala Asp Tyr Phe Ala
        115                 120                 125

gcc atc gcc aac tgc ttt acc agc aaa gac acc ccg gca ggc tgc ttc    432
Ala Ile Ala Asn Cys Phe Thr Ser Lys Asp Thr Pro Ala Gly Cys Phe
```

```
    130                 135                 140

atg atc aac aac tgc acc acc ctc tcc cca gat tca gga gat atc gcc    480
Met Ile Asn Asn Cys Thr Thr Leu Ser Pro Asp Ser Gly Asp Ile Ala
145                 150                 155                 160

aat acg ttg aaa tca cgc cat gcg atg caa gag cgc act ttg cag cag    528
Asn Thr Leu Lys Ser Arg His Ala Met Gln Glu Arg Thr Leu Gln Gln
                165                 170                 175

ttt tta tgt caa cga caa gcg cgc ggg gaa atc ccg ccc cac tgt gac    576
Phe Leu Cys Gln Arg Gln Ala Arg Gly Glu Ile Pro Pro His Cys Asp
            180                 185                 190

gtg aca cat ctg gca gaa ttc ctt aat tgt att att cag ggg atg tcg    624
Val Thr His Leu Ala Glu Phe Leu Asn Cys Ile Ile Gln Gly Met Ser
        195                 200                 205

atc agc gca cgc gaa ggt gca tcg ctg gaa aaa ctg atg cag att gcc    672
Ile Ser Ala Arg Glu Gly Ala Ser Leu Glu Lys Leu Met Gln Ile Ala
    210                 215                 220

gga acg act ttg cgt tta tgg ccc gaa ctg gtg aaa taa                711
Gly Thr Thr Leu Arg Leu Trp Pro Glu Leu Val Lys  *
225                 230                 235


<210> 189
<211> 546
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(546)

<400> 189
gtg cag gcc aaa att gcg gca tca aat acg ggt gaa ctg gat gcc ctg    48
Met Gln Ala Lys Ile Ala Ala Ser Asn Thr Gly Glu Leu Asp Ala Leu
 1               5                  10                  15

caa cag ctg gga ttc tcc ctg gta gaa ggt gaa gtt gat ttg gcg cta    96
Gln Gln Leu Gly Phe Ser Leu Val Glu Gly Glu Val Asp Leu Ala Leu
            20                  25                  30

ccc gtg aac aat gcc agt gat agc ggt gct gta gtg gca caa gag acc    144
Pro Val Asn Asn Ala Ser Asp Ser Gly Ala Val Val Ala Gln Glu Thr
        35                  40                  45

gat att ccc gca tta cgt cag tta gcc agc gcc gca ttt gcg caa agc    192
Asp Ile Pro Ala Leu Arg Gln Leu Ala Ser Ala Ala Phe Ala Gln Ser
    50                  55                  60

cgt ttt cgt gcg ccg tgg tat gcg cct gac gcc agc agt cgc ttt tat    240
Arg Phe Arg Ala Pro Trp Tyr Ala Pro Asp Ala Ser Ser Arg Phe Tyr
65                  70                  75                  80

gca cag tgg att gaa aat gcc gtg cgc ggc acc ttt gat cat caa tgt    288
Ala Gln Trp Ile Glu Asn Ala Val Arg Gly Thr Phe Asp His Gln Cys
                85                  90                  95
```

ctg att tta cgt gcg gcg tcc ggc gat att cgc ggc tat gtc tct tta 336
Leu Ile Leu Arg Ala Ala Ser Gly Asp Ile Arg Gly Tyr Val Ser Leu
100 105 110

cgg gaa ctc aat gcg aca gat gcg cga att ggc ctg ctg gct gga cgc 384
Arg Glu Leu Asn Ala Thr Asp Ala Arg Ile Gly Leu Leu Ala Gly Arg
115 120 125

ggt gca ggt gct gag ctg atg caa acg gcg cta aac tgg gcg tat cgt 432
Gly Ala Gly Ala Glu Leu Met Gln Thr Ala Leu Asn Trp Ala Tyr Arg
130 135 140

cgc ggt aaa aca act ttg cgg gtg gcg acc caa atg ggc aac acc gcc 480
Arg Gly Lys Thr Thr Leu Arg Val Ala Thr Gln Met Gly Asn Thr Ala
145 150 155 160

gcg ctt aaa cga tac ata caa agt ggt gcg aat gta gaa agc acc gcg 528
Ala Leu Lys Arg Tyr Ile Gln Ser Gly Ala Asn Val Glu Ser Thr Ala
165 170 175

tac tgg tta tac agg tga 546
Tyr Trp Leu Tyr Arg *
180

<210> 190
<211> 1131
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1131)

<400> 190
atg att cca ttt aac gca ccg ccg gtg gtg gga acc gaa ctc gac tat 48
Met Ile Pro Phe Asn Ala Pro Pro Val Val Gly Thr Glu Leu Asp Tyr
1 5 10 15

atg cag tcg gca atg ggt agc ggc aaa ctg tgt ggc gat ggc ggt ttt 96
Met Gln Ser Ala Met Gly Ser Gly Lys Leu Cys Gly Asp Gly Gly Phe
20 25 30

acc cgt cgc tgc cag cag tgg ctg gag caa cgt ttt ggc agc gcc aaa 144
Thr Arg Arg Cys Gln Gln Trp Leu Glu Gln Arg Phe Gly Ser Ala Lys
35 40 45

gtg tta ctg acg ccg tcc tgc acc gct tcg ctg gag atg gcg gcg ctg 192
Val Leu Leu Thr Pro Ser Cys Thr Ala Ser Leu Glu Met Ala Ala Leu
50 55 60

ctg ctc gat atc cag cct ggc gat gaa gtg atc atg ccg agc tac acc 240
Leu Leu Asp Ile Gln Pro Gly Asp Glu Val Ile Met Pro Ser Tyr Thr
65 70 75 80

ttt gtc tcc acc gcc aat gcc ttt gtg ctg cgt ggc gca aaa atc gtt 288
Phe Val Ser Thr Ala Asn Ala Phe Val Leu Arg Gly Ala Lys Ile Val
85 90 95

ttt gtg gat gtt cgc ccg gac acc atg aac atc gac gaa acg ctg att 336
Phe Val Asp Val Arg Pro Asp Thr Met Asn Ile Asp Glu Thr Leu Ile

```
            100                       105                      110

gaa gcg gcg atc acc gac aaa acg cgc gtt atc gtg ccg gtc cat tac      384
Glu Ala Ala Ile Thr Asp Lys Thr Arg Val Ile Val Pro Val His Tyr
        115                 120                 125

gcg ggt gtg gcc tgc gaa atg gac acc att atg gcg ttg gcg aaa aag      432
Ala Gly Val Ala Cys Glu Met Asp Thr Ile Met Ala Leu Ala Lys Lys
    130                 135                 140

cat aat ttg ttt gtg gta gaa gat gcc gct cag ggc gtg atg tcc act      480
His Asn Leu Phe Val Val Glu Asp Ala Ala Gln Gly Val Met Ser Thr
145                 150                 155                 160

tac aaa ggg cgt gca ctg gga acc att ggt cat att ggc tgc ttt agc      528
Tyr Lys Gly Arg Ala Leu Gly Thr Ile Gly His Ile Gly Cys Phe Ser
                165                 170                 175

ttc cat gaa acc aaa aac tac acg gcg ggc ggt gaa ggc ggc gcg acg      576
Phe His Glu Thr Lys Asn Tyr Thr Ala Gly Gly Glu Gly Gly Ala Thr
            180                 185                 190

ctg att aac gat aaa gcg tta atc gaa cga gcc gag atc atc cgt gaa      624
Leu Ile Asn Asp Lys Ala Leu Ile Glu Arg Ala Glu Ile Ile Arg Glu
        195                 200                 205

aag ggc act aac cgc agc cag ttc ttc cgt ggt cag gtc gat aaa tat      672
Lys Gly Thr Asn Arg Ser Gln Phe Phe Arg Gly Gln Val Asp Lys Tyr
    210                 215                 220

acc tgg cgc gat att ggc tcc agc tat ttg atg tcc gat ctg caa gct      720
Thr Trp Arg Asp Ile Gly Ser Ser Tyr Leu Met Ser Asp Leu Gln Ala
225                 230                 235                 240

gca tac ctg tgg gcg caa ctg gaa gca gcg gat cgt atc aac cag caa      768
Ala Tyr Leu Trp Ala Gln Leu Glu Ala Ala Asp Arg Ile Asn Gln Gln
                245                 250                 255

cgt ctg gcg ctg tgg caa aac tac tac gat gcg tta gcg cct ctg gcg      816
Arg Leu Ala Leu Trp Gln Asn Tyr Tyr Asp Ala Leu Ala Pro Leu Ala
            260                 265                 270

aaa gcc ggg cgt atc gag ctg ccg tcg att ccc gat ggc tgc gtg cag      864
Lys Ala Gly Arg Ile Glu Leu Pro Ser Ile Pro Asp Gly Cys Val Gln
        275                 280                 285

aac gcg cat atg ttc tac att aaa ctg cgg gat att gat gac cgg agc      912
Asn Ala His Met Phe Tyr Ile Lys Leu Arg Asp Ile Asp Asp Arg Ser
    290                 295                 300

gcg ttg att aac ttt ctg aaa gaa gcg gaa atc atg gcg gtg ttt cat      960
Ala Leu Ile Asn Phe Leu Lys Glu Ala Glu Ile Met Ala Val Phe His
305                 310                 315                 320

tac att ccg ctg cac ggt tgc cct gcg ggg gaa cac ttt ggt gag ttc      1008
Tyr Ile Pro Leu His Gly Cys Pro Ala Gly Glu His Phe Gly Glu Phe
                325                 330                 335

cac ggt gaa gat cgc tac acc acc aaa gag agc gag cgc ctg ctg cgc      1056
His Gly Glu Asp Arg Tyr Thr Thr Lys Glu Ser Glu Arg Leu Leu Arg
            340                 345                 350
```

ctg ccg ctg ttc tac aac ctg tcg ccc gtc aat cag cgt acg gta att 1104
Leu Pro Leu Phe Tyr Asn Leu Ser Pro Val Asn Gln Arg Thr Val Ile
355 360 365

gcg act ttg ttg aac tac ttt tcc tga 1131
Ala Thr Leu Leu Asn Tyr Phe Ser *
370 375

<210> 191
<211> 1251
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1251)

<400> 191
atg tcg ttg gca aaa gcg tcc ttg tgg acg gcg gcc agt aca ctg gtc 48
Met Ser Leu Ala Lys Ala Ser Leu Trp Thr Ala Ala Ser Thr Leu Val
1 5 10 15

aag att ggt gcc ggg tta ctg gtc ggt aag ttg ctg gcg gtg tca ttt 96
Lys Ile Gly Ala Gly Leu Leu Val Gly Lys Leu Leu Ala Val Ser Phe
20 25 30

ggt ccg gcg ggg ctt ggg ctg gcg gca aat ttc cgc cag ttg att acc 144
Gly Pro Ala Gly Leu Gly Leu Ala Ala Asn Phe Arg Gln Leu Ile Thr
35 40 45

gtg ctc ggc gtg ctt gcc ggg gct ggc atc ttt aac ggt gta acc aaa 192
Val Leu Gly Val Leu Ala Gly Ala Gly Ile Phe Asn Gly Val Thr Lys
50 55 60

tac gtt gcc cag tac cat gat aat ccg caa cag ctg cgc cgc gtg gtc 240
Tyr Val Ala Gln Tyr His Asp Asn Pro Gln Gln Leu Arg Arg Val Val
65 70 75 80

ggc act tca tca gcg atg gta ctt ggt ttc tct acg ctg atg gcg ctg 288
Gly Thr Ser Ser Ala Met Val Leu Gly Phe Ser Thr Leu Met Ala Leu
85 90 95

gtt ttt gtg ctg gca gct gcg cca atc agc cag gga ttg ttt ggt aat 336
Val Phe Val Leu Ala Ala Ala Pro Ile Ser Gln Gly Leu Phe Gly Asn
100 105 110

acc gac tat cag ggg ctg gtg cgt tta gtg gcg ctg gtg caa atg ggg 384
Thr Asp Tyr Gln Gly Leu Val Arg Leu Val Ala Leu Val Gln Met Gly
115 120 125

atc gcc tgg ggc aac ctg tta ctg gcg ctg atg aaa ggc ttt cgc gat 432
Ile Ala Trp Gly Asn Leu Leu Leu Ala Leu Met Lys Gly Phe Arg Asp
130 135 140

gcc gca ggt aat gcg tta tcg ctg att gtc ggc agc ttg att ggc gtt 480
Ala Ala Gly Asn Ala Leu Ser Leu Ile Val Gly Ser Leu Ile Gly Val
145 150 155 160

ctc gcg tac tac gtc agt tac cgt ttg ggc ggt tat gaa ggg gcg ttg 528

Leu Ala Tyr Tyr Val Ser Tyr Arg Leu Gly Gly Tyr Glu Gly Ala Leu
165 170 175

ctg ggt ctg gcg ctg att ccc gcg ctg gtg gta att cct gcc gcc atc 576
Leu Gly Leu Ala Leu Ile Pro Ala Leu Val Val Ile Pro Ala Ala Ile
180 185 190

atg ttg atc aaa cgt ggt gtc atc ccg tta agc tat ctg aaa ccc agc 624
Met Leu Ile Lys Arg Gly Val Ile Pro Leu Ser Tyr Leu Lys Pro Ser
195 200 205

tgg gat aac ggt ctg gca ggg cag ttg agc aaa ttt acg ctc atg gcg 672
Trp Asp Asn Gly Leu Ala Gly Gln Leu Ser Lys Phe Thr Leu Met Ala
210 215 220

ttg att acg tcg gtg acc ttg cct gtt gct tac atc atg atg cgt aaa 720
Leu Ile Thr Ser Val Thr Leu Pro Val Ala Tyr Ile Met Met Arg Lys
225 230 235 240

ctg ctg gcg gcg cag tat agc tgg gat gag gtg ggg atc tgg caa ggg 768
Leu Leu Ala Ala Gln Tyr Ser Trp Asp Glu Val Gly Ile Trp Gln Gly
245 250 255

gtg agc agt att tcc gat gcc tac ctg caa ttt att acg gca tcg ttc 816
Val Ser Ser Ile Ser Asp Ala Tyr Leu Gln Phe Ile Thr Ala Ser Phe
260 265 270

agc gta tat ttg ctg ccc acg ttg tcg cgg cta acg gaa aag cgc gat 864
Ser Val Tyr Leu Leu Pro Thr Leu Ser Arg Leu Thr Glu Lys Arg Asp
275 280 285

atc acc cgg gaa gtg gtt aaa tcg ctg aaa ttc gtc tta ccg gca gtg 912
Ile Thr Arg Glu Val Val Lys Ser Leu Lys Phe Val Leu Pro Ala Val
290 295 300

gcg gcg gcg agt ttt acc gtc tgg ctg ctg cgt gat ttt gct atc tgg 960
Ala Ala Ala Ser Phe Thr Val Trp Leu Leu Arg Asp Phe Ala Ile Trp
305 310 315 320

ctg ctg ttg tcg aat aaa ttt acc gct atg cgc gat ctc ttt gcc tgg 1008
Leu Leu Leu Ser Asn Lys Phe Thr Ala Met Arg Asp Leu Phe Ala Trp
325 330 335

cag tta gtg ggt gat gtg tta aaa gtg ggc gct tat gtc ttt ggt tat 1056
Gln Leu Val Gly Asp Val Leu Lys Val Gly Ala Tyr Val Phe Gly Tyr
340 345 350

ctg gtg atc gcc aaa gcg tca ctg cgg ttt tat att ctg gcg gaa gtc 1104
Leu Val Ile Ala Lys Ala Ser Leu Arg Phe Tyr Ile Leu Ala Glu Val
355 360 365

agc cag ttc act tta ttg atg gta ttt gcc cac tgg cta atc cct gcg 1152
Ser Gln Phe Thr Leu Leu Met Val Phe Ala His Trp Leu Ile Pro Ala
370 375 380

cat ggt gca ctg ggc gcg gcg cag gca tat atg gca act tat atc gtc 1200
His Gly Ala Leu Gly Ala Ala Gln Ala Tyr Met Ala Thr Tyr Ile Val
385 390 395 400

tat ttt tct ctt tgt tgt ggc gtg ttt tta ctc tgg cgt agg cgg gca 1248
Tyr Phe Ser Leu Cys Cys Gly Val Phe Leu Leu Trp Arg Arg Arg Ala

405 410 415

tga 1251
*

<210> 192
<211> 1353
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1353)

<400> 192
atg agt ctg ctg caa ttc agt ggc ctg ttt gtt gtc tgg ctg ctc tgc 48
Met Ser Leu Leu Gln Phe Ser Gly Leu Phe Val Val Trp Leu Leu Cys
1 5 10 15

acg ctg ttt att gcc acg ctg acc tgg ttt gag ttt cgc cgt gtg cgc 96
Thr Leu Phe Ile Ala Thr Leu Thr Trp Phe Glu Phe Arg Arg Val Arg
20 25 30

ttt aac ttc aat gtc ttc ttt tca ttg ctg ttt ttg ctc acc ttt ttc 144
Phe Asn Phe Asn Val Phe Phe Ser Leu Leu Phe Leu Leu Thr Phe Phe
35 40 45

ttc ggc ttc ccg ctg acc agc gtg ctg gta ttt cgc ttt gat gtt ggt 192
Phe Gly Phe Pro Leu Thr Ser Val Leu Val Phe Arg Phe Asp Val Gly
50 55 60

gtc gcg ccg cca gaa atc ttg ttg cag gcg ttg ctt tct gcg ggc tgc 240
Val Ala Pro Pro Glu Ile Leu Leu Gln Ala Leu Leu Ser Ala Gly Cys
65 70 75 80

ttc tac gcg gtt tac tat gtc acc tac aaa acc cgc cta cgc aaa cgc 288
Phe Tyr Ala Val Tyr Tyr Val Thr Tyr Lys Thr Arg Leu Arg Lys Arg
85 90 95

gtt gct gat gta ccg cgc cgt ccg ctg ttt acc atg aac cgc gtg gag 336
Val Ala Asp Val Pro Arg Arg Pro Leu Phe Thr Met Asn Arg Val Glu
100 105 110

acc aat ctt acg tgg gtg atc ctg atg ggt atc gcg ctg gta agc gtc 384
Thr Asn Leu Thr Trp Val Ile Leu Met Gly Ile Ala Leu Val Ser Val
115 120 125

ggc atc ttc ttc atg cac aac ggc ttt ttg ctg ttc cgg ctt aac tcc 432
Gly Ile Phe Phe Met His Asn Gly Phe Leu Leu Phe Arg Leu Asn Ser
130 135 140

tac agt cag atc ttt tcc agt gaa gtc tcc ggc gtg gcg tta aaa cgc 480
Tyr Ser Gln Ile Phe Ser Ser Glu Val Ser Gly Val Ala Leu Lys Arg
145 150 155 160

ttc ttt tac ttt ttc atc ccg gcg atg ctg gtg gtc tac ttt ctg cgc 528
Phe Phe Tyr Phe Phe Ile Pro Ala Met Leu Val Val Tyr Phe Leu Arg
165 170 175

```
cag gac agc aaa gcg tgg ctg ttt ttc ctc gtc agc acg gtc gcc ttt    576
Gln Asp Ser Lys Ala Trp Leu Phe Phe Leu Val Ser Thr Val Ala Phe
            180                 185                 190

ggc ttg ctg act tat atg att gtc ggc ggc act cgc gcc aat atc atc    624
Gly Leu Leu Thr Tyr Met Ile Val Gly Gly Thr Arg Ala Asn Ile Ile
        195                 200                 205

atc gca ttc gct atc ttc ctg ttt att ggc att att cgc ggc tgg att    672
Ile Ala Phe Ala Ile Phe Leu Phe Ile Gly Ile Ile Arg Gly Trp Ile
    210                 215                 220

tcg ttg tgg atg ctg gcg gcg gcg ggc gtg ctg ggg att gtt ggc atg    720
Ser Leu Trp Met Leu Ala Ala Ala Gly Val Leu Gly Ile Val Gly Met
225                 230                 235                 240

ttc tgg ctg gca cta aaa cgc tat gga atg aat gtg agc ggc gat gaa    768
Phe Trp Leu Ala Leu Lys Arg Tyr Gly Met Asn Val Ser Gly Asp Glu
                245                 250                 255

gcg ttc tat acg ttt ctc tat ctc act cgc gac acc ttc tcg ccg tgg    816
Ala Phe Tyr Thr Phe Leu Tyr Leu Thr Arg Asp Thr Phe Ser Pro Trp
            260                 265                 270

gag aat ctg gcg ttg ctg ttg cag aac tac gac aac atc gac ttc cag    864
Glu Asn Leu Ala Leu Leu Leu Gln Asn Tyr Asp Asn Ile Asp Phe Gln
        275                 280                 285

ggc ctg gct cca att gtc cgc gat ttc tat gtc ttt atc cct tcc tgg    912
Gly Leu Ala Pro Ile Val Arg Asp Phe Tyr Val Phe Ile Pro Ser Trp
    290                 295                 300

ctg tgg ccg ggt cgc ccg agt atg gtg ctg aac tca gcc aac tac ttt    960
Leu Trp Pro Gly Arg Pro Ser Met Val Leu Asn Ser Ala Asn Tyr Phe
305                 310                 315                 320

acc tgg gaa gtg ctg aat aac cac tcc gga ctg gcg atc tcg cct acg    1008
Thr Trp Glu Val Leu Asn Asn His Ser Gly Leu Ala Ile Ser Pro Thr
                325                 330                 335

ctt ata ggc tca ctg gtg gtg atg ggc ggc gcg ttg ttc atc ccg ctc    1056
Leu Ile Gly Ser Leu Val Val Met Gly Gly Ala Leu Phe Ile Pro Leu
            340                 345                 350

ggg gcg atc gtg gtt ggt ctg atc atc aaa tgg ttc gac tgg ctg tat    1104
Gly Ala Ile Val Val Gly Leu Ile Ile Lys Trp Phe Asp Trp Leu Tyr
        355                 360                 365

gag ctg ggc aac cgc gag cct aat cgc tat aaa gct gcg ata ttg cac    1152
Glu Leu Gly Asn Arg Glu Pro Asn Arg Tyr Lys Ala Ala Ile Leu His
    370                 375                 380

agt ttc tgc ttt ggg gcg atc ttc aat atg atc gtg ctg gcg cgt gaa    1200
Ser Phe Cys Phe Gly Ala Ile Phe Asn Met Ile Val Leu Ala Arg Glu
385                 390                 395                 400

ggg ctg gat tcg ttt gtc tca cgc gtg gtc ttt ttt atc gtg gtc ttc    1248
Gly Leu Asp Ser Phe Val Ser Arg Val Val Phe Phe Ile Val Val Phe
                405                 410                 415

ggc gca tgt ctg atg atc gca aaa ctg ttg tac tgg ctt ttt gaa agc    1296
```

Gly Ala Cys Leu Met Ile Ala Lys Leu Leu Tyr Trp Leu Phe Glu Ser
420 425 430

gcc gga ctc att cat aaa cgt aca aaa tca tcg ctc cgg acg cag gtt 1344
Ala Gly Leu Ile His Lys Arg Thr Lys Ser Ser Leu Arg Thr Gln Val
435 440 445

gaa gga taa 1353
Glu Gly *
450

<210> 193
<211> 741
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(741)

<400> 193
atg aat aac aac acc acg gca cca acc tat acg ctg cgt ggc tta cag 48
Met Asn Asn Asn Thr Thr Ala Pro Thr Tyr Thr Leu Arg Gly Leu Gln
1 5 10 15

ttg att ggt tgg cgt gat atg cag cac gcc ctc gat tat ctg ttt gct 96
Leu Ile Gly Trp Arg Asp Met Gln His Ala Leu Asp Tyr Leu Phe Ala
20 25 30

gac ggg cag ctt aag cag gga acg ctg gtt gcc att aat gct gaa aaa 144
Asp Gly Gln Leu Lys Gln Gly Thr Leu Val Ala Ile Asn Ala Glu Lys
35 40 45

atg ctg act att gaa gat aac gcc gag gtc agg gag tta att aac gct 192
Met Leu Thr Ile Glu Asp Asn Ala Glu Val Arg Glu Leu Ile Asn Ala
50 55 60

gcc gaa ttt aaa tat gcg gat ggc atc agc gtt gta cgt tca gta cgt 240
Ala Glu Phe Lys Tyr Ala Asp Gly Ile Ser Val Val Arg Ser Val Arg
65 70 75 80

aaa aag tac ccg cag gcg cag gtt tcc cgc gtt gcc ggt gcc gat ctc 288
Lys Lys Tyr Pro Gln Ala Gln Val Ser Arg Val Ala Gly Ala Asp Leu
85 90 95

tgg gaa gag ctg atg gcg cgc gca ggc aaa gaa ggg acg ccg gta ttt 336
Trp Glu Glu Leu Met Ala Arg Ala Gly Lys Glu Gly Thr Pro Val Phe
100 105 110

ctt gtg ggc ggt aaa cct gaa gtg ctg gcg caa act gaa gct aaa ctg 384
Leu Val Gly Gly Lys Pro Glu Val Leu Ala Gln Thr Glu Ala Lys Leu
115 120 125

cgc aac cag tgg aat gtg aat atc gtt ggc agt cag gat ggt tat ttt 432
Arg Asn Gln Trp Asn Val Asn Ile Val Gly Ser Gln Asp Gly Tyr Phe
130 135 140

aaa ccc gag cag cgt cag gcg ctg ttt gaa cgc att cat gcc agc ggt 480
Lys Pro Glu Gln Arg Gln Ala Leu Phe Glu Arg Ile His Ala Ser Gly
145 150 155 160

```
gcg caa atc gtc acc gtt gcg atg gga tcg cca aag cag gag atc atc    528
Ala Gln Ile Val Thr Val Ala Met Gly Ser Pro Lys Gln Glu Ile Ile
                165                 170                 175

atg cgc gac tgc cgt ctg gta cat cca gat gcg ctg tat atg ggc gtt    576
Met Arg Asp Cys Arg Leu Val His Pro Asp Ala Leu Tyr Met Gly Val
            180                 185                 190

ggc ggg act tac gat gtt ttc acc ggt cac gta aaa cgc gca ccg aaa    624
Gly Gly Thr Tyr Asp Val Phe Thr Gly His Val Lys Arg Ala Pro Lys
        195                 200                 205

atc tgg caa acg ctg ggg ctg gag tgg ctc tac cgc ctg ctt tcg cag    672
Ile Trp Gln Thr Leu Gly Leu Glu Trp Leu Tyr Arg Leu Leu Ser Gln
    210                 215                 220

ccg agc cgc att aag cgt cag ctt cgt ttg ctg cgt tat tta cgc tgg    720
Pro Ser Arg Ile Lys Arg Gln Leu Arg Leu Leu Arg Tyr Leu Arg Trp
225                 230                 235                 240

cac tac acc ggc aac cta tga                                        741
His Tyr Thr Gly Asn Leu  *
                245
```

<210> 194
<211> 225
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(225)

<400> 194

```
atg act gta ctg att cac gta ctg gga tcg gat atc cct cac cat aac    48
Met Thr Val Leu Ile His Val Leu Gly Ser Asp Ile Pro His His Asn
 1               5                  10                  15

cga acc gtt ttg cgg ttt ttc aat gac gcg ctg gcc gcg acg agc gag    96
Arg Thr Val Leu Arg Phe Phe Asn Asp Ala Leu Ala Ala Thr Ser Glu
            20                  25                  30

cac gcg cgc gag ttt atg gtt gtt ggc aag gac gac ggc tta agt gat    144
His Ala Arg Glu Phe Met Val Val Gly Lys Asp Asp Gly Leu Ser Asp
        35                  40                  45

agc tgt ccg gcg ctt tct gtg caa ttt ttc ccc tgg gaa aaa atc gct    192
Ser Cys Pro Ala Leu Ser Val Gln Phe Phe Pro Trp Glu Lys Ile Ala
    50                  55                  60

ggc gga agc ggt cat cgc gaa agc aaa agc taa                        225
Gly Gly Ser Gly His Arg Glu Ser Lys Ser  *
65                  70
```

<210> 195
<211> 615
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(615)

<400> 195
gtg cgg ggc gaa ctg ctg ttc ttc ccg acg cgg atg gac cct tcg ctc 48
Met Arg Gly Glu Leu Leu Phe Phe Pro Thr Arg Met Asp Pro Ser Leu
1 5 10 15

aat acg atg gcg aac gat cgg caa cgt gaa ggg aaa atg acc att ctg 96
Asn Thr Met Ala Asn Asp Arg Gln Arg Glu Gly Lys Met Thr Ile Leu
20 25 30

gtg ggg aac tcc ggc gac cgc agc aat gag cat att gct gcc ttg cgc 144
Val Gly Asn Ser Gly Asp Arg Ser Asn Glu His Ile Ala Ala Leu Arg
35 40 45

gcc gtt cat cag caa ttt ggc gat acg gta aaa gtg gtg gtg ccg atg 192
Ala Val His Gln Gln Phe Gly Asp Thr Val Lys Val Val Val Pro Met
50 55 60

gga tat ccg cct aat aac gaa gcg tac att gag gaa gtt cgt cag gcg 240
Gly Tyr Pro Pro Asn Asn Glu Ala Tyr Ile Glu Glu Val Arg Gln Ala
65 70 75 80

ggg ctg gag tta ttc agc gaa gaa aat cta caa att ctg agc gaa aaa 288
Gly Leu Glu Leu Phe Ser Glu Glu Asn Leu Gln Ile Leu Ser Glu Lys
85 90 95

ctg gaa ttt gac gcc tat ctg gcg cta ctt cgt cag tgc gat ctt ggt 336
Leu Glu Phe Asp Ala Tyr Leu Ala Leu Leu Arg Gln Cys Asp Leu Gly
100 105 110

tac ttt att ttt gcc cgc cag cag ggc att ggt acg ctg tgc tta ctg 384
Tyr Phe Ile Phe Ala Arg Gln Gln Gly Ile Gly Thr Leu Cys Leu Leu
115 120 125

att cag gcg ggc att cct tgt gtg ctt aac cgg gaa aat ccg ttc tgg 432
Ile Gln Ala Gly Ile Pro Cys Val Leu Asn Arg Glu Asn Pro Phe Trp
130 135 140

cag gat atg acg gaa caa cat ttg ccg gtg ctg ttt act acc gac gat 480
Gln Asp Met Thr Glu Gln His Leu Pro Val Leu Phe Thr Thr Asp Asp
145 150 155 160

ctc aac gag gat att gtg cgt gaa gcg cag cgc cag ttg gcg tcg gtg 528
Leu Asn Glu Asp Ile Val Arg Glu Ala Gln Arg Gln Leu Ala Ser Val
165 170 175

gat aaa aac acc att gcc ttc ttt agc cct aac tat cta caa ggc tgg 576
Asp Lys Asn Thr Ile Ala Phe Phe Ser Pro Asn Tyr Leu Gln Gly Trp
180 185 190

cag cgg gcg ttg gcg att gcc gcc agg gag gtc gca tga 615
Gln Arg Ala Leu Ala Ile Ala Ala Arg Glu Val Ala *
195 200

<210> 196
<211> 549

<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(549)

<400> 196
atg att cgg caa cgt cgt cgt gcg tta acg ccg gaa caa cag cag gaa 48
Met Ile Arg Gln Arg Arg Arg Ala Leu Thr Pro Glu Gln Gln Gln Glu
1 5 10 15

atg ggt caa caa gcc gct acc cgg atg atg act tat ccc ccg gtg gtg 96
Met Gly Gln Gln Ala Ala Thr Arg Met Met Thr Tyr Pro Pro Val Val
20 25 30

atg gca cat acg gtc gct gta ttc ctc tct ttt gat ggc gaa ctc gac 144
Met Ala His Thr Val Ala Val Phe Leu Ser Phe Asp Gly Glu Leu Asp
35 40 45

acc cag cca ctc ata gaa caa ctc tgg cgc gcc ggt aag cgc gta tat 192
Thr Gln Pro Leu Ile Glu Gln Leu Trp Arg Ala Gly Lys Arg Val Tyr
50 55 60

ctt cca gtt ttg cat ccc ttt agt gcc ggt aat ttg ctg ttc ctg aat 240
Leu Pro Val Leu His Pro Phe Ser Ala Gly Asn Leu Leu Phe Leu Asn
65 70 75 80

tac cat ccg caa agc gaa ctg gtg atg aac agg ttg aag atc cat gag 288
Tyr His Pro Gln Ser Glu Leu Val Met Asn Arg Leu Lys Ile His Glu
85 90 95

cca aaa ttg gat gtg cgt gac gtg cta ccc ctt tcc cga tta gac gtg 336
Pro Lys Leu Asp Val Arg Asp Val Leu Pro Leu Ser Arg Leu Asp Val
100 105 110

ctg atc aca ccg ctg gtc gcc ttt gat gag tac ggt cag cgc ctg gga 384
Leu Ile Thr Pro Leu Val Ala Phe Asp Glu Tyr Gly Gln Arg Leu Gly
115 120 125

atg ggc ggt ggt ttt tat gat cgg acc tta caa aac tgg cag cac tat 432
Met Gly Gly Gly Phe Tyr Asp Arg Thr Leu Gln Asn Trp Gln His Tyr
130 135 140

aaa acg caa ccg gtg ggt tat gcg cat gat tgt cag ttg gtg gaa aaa 480
Lys Thr Gln Pro Val Gly Tyr Ala His Asp Cys Gln Leu Val Glu Lys
145 150 155 160

ctc ccc gtt gaa gag tgg gat atc cct ctt cct gcg gtg gtt aca ccg 528
Leu Pro Val Glu Glu Trp Asp Ile Pro Leu Pro Ala Val Val Thr Pro
165 170 175

tcg aaa gtc tgg gag tgg taa 549
Ser Lys Val Trp Glu Trp *
180

<210> 197
<211> 276
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(276)

<400> 197
atg gca cgc gta act gtt cag gac gct gta gag aaa att ggt aac cgt 48
Met Ala Arg Val Thr Val Gln Asp Ala Val Glu Lys Ile Gly Asn Arg
1 5 10 15

ttt gac ctg gta ctg gtc gcc gcg cgt cgc gct cgt cag atg cag gta 96
Phe Asp Leu Val Leu Val Ala Ala Arg Arg Ala Arg Gln Met Gln Val
20 25 30

ggc gga aag gat ccg ctg gta ccg gaa gaa aac gat aaa acc act gta 144
Gly Gly Lys Asp Pro Leu Val Pro Glu Glu Asn Asp Lys Thr Thr Val
35 40 45

atc gcg ctg cgc gaa atc gaa gaa ggt ctg atc aac aac cag atc ctc 192
Ile Ala Leu Arg Glu Ile Glu Glu Gly Leu Ile Asn Asn Gln Ile Leu
50 55 60

gac gtt cgc gaa cgc cag gaa cag caa gag cag gaa gcc gct gaa tta 240
Asp Val Arg Glu Arg Gln Glu Gln Gln Glu Gln Glu Ala Ala Glu Leu
65 70 75 80

caa gcc gtt acc gct att gct gaa ggt cgt cgt taa 276
Gln Ala Val Thr Ala Ile Ala Glu Gly Arg Arg *
85 90

<210> 198
<211> 2109
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(2109)

<400> 198
ttg tat ctg ttt gaa agc ctg aat caa ctg att caa acc tac ctg ccg 48
Met Tyr Leu Phe Glu Ser Leu Asn Gln Leu Ile Gln Thr Tyr Leu Pro
1 5 10 15

gaa gac caa atc aag cgt ctg cgg cag gcg tat ctc gtt gca cgt gat 96
Glu Asp Gln Ile Lys Arg Leu Arg Gln Ala Tyr Leu Val Ala Arg Asp
20 25 30

gct cac gag ggg caa aca cgt tca agc ggt gaa ccc tat atc acg cac 144
Ala His Glu Gly Gln Thr Arg Ser Ser Gly Glu Pro Tyr Ile Thr His
35 40 45

ccg gta gcg gtt gcc tgc att ctg gcc gag atg aaa ctc gac tat gaa 192
Pro Val Ala Val Ala Cys Ile Leu Ala Glu Met Lys Leu Asp Tyr Glu
50 55 60

acg ctg atg gcg gcg ctg ctg cat gac gtg att gaa gat act ccc gcc 240
Thr Leu Met Ala Ala Leu Leu His Asp Val Ile Glu Asp Thr Pro Ala
65 70 75 80

acc tac cag gat atg gaa cag ctt ttt ggt aaa agc gtc gcc gag ctg 288
Thr Tyr Gln Asp Met Glu Gln Leu Phe Gly Lys Ser Val Ala Glu Leu
85 90 95

gta gag ggg gtg tcg aaa ctt gat aaa ctc aag ttc cgc gat aag aaa 336
Val Glu Gly Val Ser Lys Leu Asp Lys Leu Lys Phe Arg Asp Lys Lys
100 105 110

gag gcg cag gcc gaa aac ttt cgc aag atg att atg gcg atg gtg cag 384
Glu Ala Gln Ala Glu Asn Phe Arg Lys Met Ile Met Ala Met Val Gln
115 120 125

gat atc cgc gtc atc ctc atc aaa ctt gcc gac cgt acc cac aac atg 432
Asp Ile Arg Val Ile Leu Ile Lys Leu Ala Asp Arg Thr His Asn Met
130 135 140

cgc acg ctg ggc tca ctt cgc ccg gac aaa cgt cgc cgc atc gcc cgt 480
Arg Thr Leu Gly Ser Leu Arg Pro Asp Lys Arg Arg Arg Ile Ala Arg
145 150 155 160

gaa act ctc gaa att tat agc ccg ctg gcg cac cgt tta ggt atc cac 528
Glu Thr Leu Glu Ile Tyr Ser Pro Leu Ala His Arg Leu Gly Ile His
165 170 175

cac att aaa acc gaa ctc gaa gag ctg ggt ttt gag gcg ctg tat ccc 576
His Ile Lys Thr Glu Leu Glu Glu Leu Gly Phe Glu Ala Leu Tyr Pro
180 185 190

aac cgt tat cgc gta atc aaa gaa gtg gtg aaa gcc gcg cgc ggc aac 624
Asn Arg Tyr Arg Val Ile Lys Glu Val Val Lys Ala Ala Arg Gly Asn
195 200 205

cgt aaa gag atg atc cag aag att ctt tct gaa atc gaa ggg cgt ttg 672
Arg Lys Glu Met Ile Gln Lys Ile Leu Ser Glu Ile Glu Gly Arg Leu
210 215 220

cag gaa gcg gga ata ccg tgc cgc gtc agt ggt cgc gag aag cat ctt 720
Gln Glu Ala Gly Ile Pro Cys Arg Val Ser Gly Arg Glu Lys His Leu
225 230 235 240

tat tcg att tac tgc aaa atg gtg ctc aaa gag cag cgt ttt cac tcg 768
Tyr Ser Ile Tyr Cys Lys Met Val Leu Lys Glu Gln Arg Phe His Ser
245 250 255

atc atg gac atc tac gct ttc cgc gtg atc gtc aat gat tct gac acc 816
Ile Met Asp Ile Tyr Ala Phe Arg Val Ile Val Asn Asp Ser Asp Thr
260 265 270

tgt tat cgc gtg ctg ggc cag atg cac agc ctg tac aag ccg cgt ccg 864
Cys Tyr Arg Val Leu Gly Gln Met His Ser Leu Tyr Lys Pro Arg Pro
275 280 285

ggc cgc gtg aaa gac tat atc gcc att cca aaa gcg aac ggc tat cag 912
Gly Arg Val Lys Asp Tyr Ile Ala Ile Pro Lys Ala Asn Gly Tyr Gln
290 295 300

tct ttg cac acc tcg atg atc ggc ccg cac ggt gtg ccg gtt gag gtc 960
Ser Leu His Thr Ser Met Ile Gly Pro His Gly Val Pro Val Glu Val
305 310 315 320

cag atc cgt acc gaa gat atg gac cag atg gcg gag atg ggt gtt gcc 1008

Gln Ile Arg Thr Glu Asp Met Asp Gln Met Ala Glu Met Gly Val Ala
325 330 335

gcg cac tgg gct tat aaa gag cac ggc gaa acc agt act acc gca caa 1056
Ala His Trp Ala Tyr Lys Glu His Gly Glu Thr Ser Thr Thr Ala Gln
340 345 350

atc cgc gcc cag cgc tgg atg caa agc ctg ctg gag ctg caa cag agc 1104
Ile Arg Ala Gln Arg Trp Met Gln Ser Leu Leu Glu Leu Gln Gln Ser
355 360 365

gcc ggt agt tcg ttt gaa ttt atc gag agc gtt aaa tcc gat ctc ttc 1152
Ala Gly Ser Ser Phe Glu Phe Ile Glu Ser Val Lys Ser Asp Leu Phe
370 375 380

ccg gat gag att tac gtt ttc aca ccg gaa ggg cgc att gtc gag ctg 1200
Pro Asp Glu Ile Tyr Val Phe Thr Pro Glu Gly Arg Ile Val Glu Leu
385 390 395 400

cct gcc ggt gca acg ccc gtc gac ttc gct tat gca gtg cat acc gat 1248
Pro Ala Gly Ala Thr Pro Val Asp Phe Ala Tyr Ala Val His Thr Asp
405 410 415

atc ggt cat gcc tgc gtg ggc gca cgc gtt gac cgc cag cct tac ccg 1296
Ile Gly His Ala Cys Val Gly Ala Arg Val Asp Arg Gln Pro Tyr Pro
420 425 430

ctg tcg cag ccg ctt acc agc ggt caa acc gtt gaa atc att acc gct 1344
Leu Ser Gln Pro Leu Thr Ser Gly Gln Thr Val Glu Ile Ile Thr Ala
435 440 445

ccg ggc gct cgc ccg aat gcc gct tgg ctg aac ttt gtc gtt agc tcg 1392
Pro Gly Ala Arg Pro Asn Ala Ala Trp Leu Asn Phe Val Val Ser Ser
450 455 460

aaa gcg cgc gcc aaa att cgt cag ttg ctg aaa aac ctc aag cgt gat 1440
Lys Ala Arg Ala Lys Ile Arg Gln Leu Leu Lys Asn Leu Lys Arg Asp
465 470 475 480

gat tct gta agc ctg ggc cgt cgt ctg ctc aac cat gct ttg ggt ggt 1488
Asp Ser Val Ser Leu Gly Arg Arg Leu Leu Asn His Ala Leu Gly Gly
485 490 495

agc cgt aag ctg aat gaa atc ccg cag gaa aat att cag cgc gag ctg 1536
Ser Arg Lys Leu Asn Glu Ile Pro Gln Glu Asn Ile Gln Arg Glu Leu
500 505 510

gat cgc atg aag ctg gca acg ctt gac gat ctg ctg gca gaa atc gga 1584
Asp Arg Met Lys Leu Ala Thr Leu Asp Asp Leu Leu Ala Glu Ile Gly
515 520 525

ctt ggt aac gca atg agc gtg gtg gtc gcg aaa aat ctg caa cat ggg 1632
Leu Gly Asn Ala Met Ser Val Val Val Ala Lys Asn Leu Gln His Gly
530 535 540

gac gcc tcc att cca ccg gca acc caa agc cac gga cat ctg ccc att 1680
Asp Ala Ser Ile Pro Pro Ala Thr Gln Ser His Gly His Leu Pro Ile
545 550 555 560

aaa ggt gcc gat ggc gtg ctg atc acc ttt gcg aaa tgc tgc cgc cct 1728
Lys Gly Ala Asp Gly Val Leu Ile Thr Phe Ala Lys Cys Cys Arg Pro

```
                565                 570                 575

att cct ggc gac ccg att atc gcc cac gtc agc ccc ggt aaa ggt ctg    1776
Ile Pro Gly Asp Pro Ile Ile Ala His Val Ser Pro Gly Lys Gly Leu
            580                 585                 590

gtg atc cac cat gaa tcc tgc cgt aat atc cgt ggc tac cag aaa gag    1824
Val Ile His His Glu Ser Cys Arg Asn Ile Arg Gly Tyr Gln Lys Glu
        595                 600                 605

cca gag aag ttt atg gct gtg gaa tgg gat aaa gag acg gcg cag gag    1872
Pro Glu Lys Phe Met Ala Val Glu Trp Asp Lys Glu Thr Ala Gln Glu
    610                 615                 620

ttc atc acc gaa atc aag gtg gag atg ttc aat cat cag ggt gcg ctg    1920
Phe Ile Thr Glu Ile Lys Val Glu Met Phe Asn His Gln Gly Ala Leu
625                 630                 635                 640

gca aac ctg acg gcg gca att aac acc acg act tcg aat att caa agt    1968
Ala Asn Leu Thr Ala Ala Ile Asn Thr Thr Thr Ser Asn Ile Gln Ser
                645                 650                 655

ttg aat acg gaa gag aaa gat ggt cgc gtc tac agc gcc ttt att cgt    2016
Leu Asn Thr Glu Glu Lys Asp Gly Arg Val Tyr Ser Ala Phe Ile Arg
            660                 665                 670

ctg acc gct cgt gac cgt gtg cat ctg gcg aat atc atg cgc aaa atc    2064
Leu Thr Ala Arg Asp Arg Val His Leu Ala Asn Ile Met Arg Lys Ile
        675                 680                 685

cgc gtg atg cca gac gtg att aaa gtc acc cga aac cga aat taa        2109
Arg Val Met Pro Asp Val Ile Lys Val Thr Arg Asn Arg Asn  *
    690                 695                 700
```

<210> 199
<211> 690
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(690)

<400> 199

```
atg aac cca aca cgt tat gca cgc atc tgc gaa atg ctc gcc agg cgg    48
Met Asn Pro Thr Arg Tyr Ala Arg Ile Cys Glu Met Leu Ala Arg Arg
 1               5                  10                  15

cag cct gat ctg acc gtc tgc atg gag cag gtc cac aaa cct cat aac    96
Gln Pro Asp Leu Thr Val Cys Met Glu Gln Val His Lys Pro His Asn
            20                  25                  30

gtt tct gcg att att cgt acc gca gat gcc gtt ggc gta cat gaa gtt    144
Val Ser Ala Ile Ile Arg Thr Ala Asp Ala Val Gly Val His Glu Val
        35                  40                  45

cac gcc gtc tgg cct ggt agc cgc atg cgc acc atg gct tcg gca gcg    192
His Ala Val Trp Pro Gly Ser Arg Met Arg Thr Met Ala Ser Ala Ala
    50                  55                  60
```

```
gcg ggt agt aac agc tgg gta cag gtg aaa aca cac cgc acc att ggc    240
Ala Gly Ser Asn Ser Trp Val Gln Val Lys Thr His Arg Thr Ile Gly
 65                  70                  75                  80

gat gcc gtc gct cat ctc aaa ggc cag ggc atg cag att ctg gca acc    288
Asp Ala Val Ala His Leu Lys Gly Gln Gly Met Gln Ile Leu Ala Thr
                 85                  90                  95

cat ctt tct gat aac gct gtc gat ttc cgc gaa att gat tac act cgc    336
His Leu Ser Asp Asn Ala Val Asp Phe Arg Glu Ile Asp Tyr Thr Arg
            100                 105                 110

ccg acc tgc att ttg atg gga cag gag aaa acg ggc atc acg cag gaa    384
Pro Thr Cys Ile Leu Met Gly Gln Glu Lys Thr Gly Ile Thr Gln Glu
        115                 120                 125

gca ttg gcc ctg gcg gat cag gac atc atc att ccg atg atc ggc atg    432
Ala Leu Ala Leu Ala Asp Gln Asp Ile Ile Ile Pro Met Ile Gly Met
    130                 135                 140

gtg cag tcg ctg aat gtt tcc gtt gcc tca gcc ctc att ctt tac gaa    480
Val Gln Ser Leu Asn Val Ser Val Ala Ser Ala Leu Ile Leu Tyr Glu
145                 150                 155                 160

gcc cag cgt cag cgg caa aat gca ggc atg tac ctg cgt gaa aac agc    528
Ala Gln Arg Gln Arg Gln Asn Ala Gly Met Tyr Leu Arg Glu Asn Ser
                165                 170                 175

atg ttg ccg gaa gca gag caa caa cgc ctg ttg ttt gaa ggc ggc tat    576
Met Leu Pro Glu Ala Glu Gln Gln Arg Leu Leu Phe Glu Gly Gly Tyr
            180                 185                 190

ccg gtg ctg gcg aaa gtc gca aaa cgc aaa ggc ctg cct tat ccc cac    624
Pro Val Leu Ala Lys Val Ala Lys Arg Lys Gly Leu Pro Tyr Pro His
        195                 200                 205

gtc aat cag caa ggc gag atc gaa gct gat gcc gac tgg tgg gct act    672
Val Asn Gln Gln Gly Glu Ile Glu Ala Asp Ala Asp Trp Trp Ala Thr
    210                 215                 220

atg cag gct gca ggg taa                                            690
Met Gln Ala Ala Gly  *
225
```

<210> 200
<211> 2082
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(2082)

<400> 200

```
atg aaa ggt cgc ctg tta gat gct gtc cca ctc agt tcc cta acg ggc    48
Met Lys Gly Arg Leu Leu Asp Ala Val Pro Leu Ser Ser Leu Thr Gly
 1               5                  10                  15

gtt ggc gca gca ctt agt aac aaa ctg gcg aaa atc aac ctg cat acc    96
```

Val Gly Ala Ala Leu Ser Asn Lys Leu Ala Lys Ile Asn Leu His Thr
20 25 30

gtg cag gat cta ctc tta cac ctt ccc ctg cgc tac gaa gat cgc acc 144
Val Gln Asp Leu Leu Leu His Leu Pro Leu Arg Tyr Glu Asp Arg Thr
35 40 45

cat ctc tac ccc atc gga gaa cta ctg ccg ggc gtt tat gcc acg gtg 192
His Leu Tyr Pro Ile Gly Glu Leu Leu Pro Gly Val Tyr Ala Thr Val
50 55 60

gaa ggc gaa gtg ctg aac tgc aat atc tcc ttc ggc ggt cgg cgg atg 240
Glu Gly Glu Val Leu Asn Cys Asn Ile Ser Phe Gly Gly Arg Arg Met
65 70 75 80

atg acc tgc cag atc agc gac ggt tcc ggc atc ctc acc atg cgc ttt 288
Met Thr Cys Gln Ile Ser Asp Gly Ser Gly Ile Leu Thr Met Arg Phe
85 90 95

ttc aat ttc agc gcg gca atg aaa aat agc ctg gcg gcg ggc cgc cgt 336
Phe Asn Phe Ser Ala Ala Met Lys Asn Ser Leu Ala Ala Gly Arg Arg
100 105 110

gta ctg gct tat ggc gaa gca aag cgc ggt aaa tat ggt gcg gag atg 384
Val Leu Ala Tyr Gly Glu Ala Lys Arg Gly Lys Tyr Gly Ala Glu Met
115 120 125

atc cac ccg gaa tac cgc gtg cag ggc gat ctc agc acg cca gaa tta 432
Ile His Pro Glu Tyr Arg Val Gln Gly Asp Leu Ser Thr Pro Glu Leu
130 135 140

cag gaa acg ctc acg ccg gtt tat cca aca acg gaa ggc gta aag cag 480
Gln Glu Thr Leu Thr Pro Val Tyr Pro Thr Thr Glu Gly Val Lys Gln
145 150 155 160

gcc acg ctg cgt aaa tta acc gac cag gcg ctg gat ctg ctc gac acc 528
Ala Thr Leu Arg Lys Leu Thr Asp Gln Ala Leu Asp Leu Leu Asp Thr
165 170 175

tgc gcc att gaa gaa ctc ctg ccg ccg gaa ctg tca caa gga atg atg 576
Cys Ala Ile Glu Glu Leu Leu Pro Pro Glu Leu Ser Gln Gly Met Met
180 185 190

acg cta ccg gaa gcg ttg cgc act ttg cac cgc ccg cca ccg acg cta 624
Thr Leu Pro Glu Ala Leu Arg Thr Leu His Arg Pro Pro Pro Thr Leu
195 200 205

cag ctt agc gat ctg gaa acc ggg cag cat ccg gcg caa cgt cgt ctg 672
Gln Leu Ser Asp Leu Glu Thr Gly Gln His Pro Ala Gln Arg Arg Leu
210 215 220

att ctg gaa gaa ctg ctg gcg cac aac ctc agc atg tta gcc tta cgt 720
Ile Leu Glu Glu Leu Leu Ala His Asn Leu Ser Met Leu Ala Leu Arg
225 230 235 240

gcc gga gca cag cgt ttt cat gcc cag ccg ctg agc gcc aat gac acg 768
Ala Gly Ala Gln Arg Phe His Ala Gln Pro Leu Ser Ala Asn Asp Thr
245 250 255

ctg aaa aat aaa ctc ctc gcc gcc tta ccg ttc aag cca acg ggc gca 816
Leu Lys Asn Lys Leu Leu Ala Ala Leu Pro Phe Lys Pro Thr Gly Ala

260 265 270

```
cag gca cgc gta gtg gcg gag atc gag cgc gat atg gcg ctg gat gtg    864
Gln Ala Arg Val Val Ala Glu Ile Glu Arg Asp Met Ala Leu Asp Val
        275                 280                 285

ccg atg atg cgt ctg gtg cag ggc gat gta ggt tcc ggt aaa acg ctg    912
Pro Met Met Arg Leu Val Gln Gly Asp Val Gly Ser Gly Lys Thr Leu
    290                 295                 300

gtc gcc gcc ctc gcc gcg ttg cgt gcg att gcc cac ggc aaa cag gta    960
Val Ala Ala Leu Ala Ala Leu Arg Ala Ile Ala His Gly Lys Gln Val
305                 310                 315                 320

gca ttg atg gca cca acc gaa tta ctc gcc gag cag cac gcc aat aac   1008
Ala Leu Met Ala Pro Thr Glu Leu Leu Ala Glu Gln His Ala Asn Asn
                325                 330                 335

ttc cgc aac tgg ttt gca ccg ctc ggt atc gaa gtg ggc tgg ctc gcc   1056
Phe Arg Asn Trp Phe Ala Pro Leu Gly Ile Glu Val Gly Trp Leu Ala
            340                 345                 350

ggt aag cag aaa ggt aaa gca cgg ctg gca cag cag gaa gcc atc gcc   1104
Gly Lys Gln Lys Gly Lys Ala Arg Leu Ala Gln Gln Glu Ala Ile Ala
        355                 360                 365

agc ggt cag gtg cag atg att gtc ggt aca cac gcc atc ttc cag gaa   1152
Ser Gly Gln Val Gln Met Ile Val Gly Thr His Ala Ile Phe Gln Glu
    370                 375                 380

cag gtg cag ttt aac ggc ctg gcg ctg gtg att atc gac gaa cag cat   1200
Gln Val Gln Phe Asn Gly Leu Ala Leu Val Ile Ile Asp Glu Gln His
385                 390                 395                 400

cgt ttt ggc gtg cat cag cgt ctg gca ttg tgg gag aaa ggc cag cag   1248
Arg Phe Gly Val His Gln Arg Leu Ala Leu Trp Glu Lys Gly Gln Gln
                405                 410                 415

cag ggc ttc cat ccg cat cag ttg atc atg acc gcc acg ccg atc ccc   1296
Gln Gly Phe His Pro His Gln Leu Ile Met Thr Ala Thr Pro Ile Pro
            420                 425                 430

cgc acg ctg gca atg act gcg tat gcc gat ctc gat acc tcg gtg ata   1344
Arg Thr Leu Ala Met Thr Ala Tyr Ala Asp Leu Asp Thr Ser Val Ile
        435                 440                 445

gat gag ctg ccg cca ggc cgc acg cca gtg act acg gtc gct att cct   1392
Asp Glu Leu Pro Pro Gly Arg Thr Pro Val Thr Thr Val Ala Ile Pro
    450                 455                 460

gat acc cgc cgt acc gac atc att gac cgc gtg cac cac gcc tgc ata   1440
Asp Thr Arg Arg Thr Asp Ile Ile Asp Arg Val His His Ala Cys Ile
465                 470                 475                 480

act gaa ggt cgt cag gca tac tgg gtt tgt acg ttg att gaa gag tcg   1488
Thr Glu Gly Arg Gln Ala Tyr Trp Val Cys Thr Leu Ile Glu Glu Ser
                485                 490                 495

gaa ttg ctg gaa gcg cag gcg gcg gaa gct acc tgg gaa gag ttg aaa   1536
Glu Leu Leu Glu Ala Gln Ala Ala Glu Ala Thr Trp Glu Glu Leu Lys
            500                 505                 510
```

ctg gcg cta cca gag ttg aac gtt ggc ctg gta cac ggg cgg atg aaa 1584
Leu Ala Leu Pro Glu Leu Asn Val Gly Leu Val His Gly Arg Met Lys
515 520 525

cct gcc gag aaa cag gcg gtg atg gcg tcg ttt aaa caa ggt gag cta 1632
Pro Ala Glu Lys Gln Ala Val Met Ala Ser Phe Lys Gln Gly Glu Leu
530 535 540

cac ctg ctg gtt gcc aca acc gtt att gaa gtc ggc gtt gat gtg cct 1680
His Leu Leu Val Ala Thr Thr Val Ile Glu Val Gly Val Asp Val Pro
545 550 555 560

aac gcc agt ctg atg att atc gaa aac ccg gag cgt ctg ggt ctg gcg 1728
Asn Ala Ser Leu Met Ile Ile Glu Asn Pro Glu Arg Leu Gly Leu Ala
565 570 575

cag tta cac cag ctg cgc ggg cgc gta ggt cgt ggc gcg gtg gct tct 1776
Gln Leu His Gln Leu Arg Gly Arg Val Gly Arg Gly Ala Val Ala Ser
580 585 590

cac tgc gtg ctg ctc tac aaa acg ccg ctt tct aaa acg gcg caa att 1824
His Cys Val Leu Leu Tyr Lys Thr Pro Leu Ser Lys Thr Ala Gln Ile
595 600 605

cgc ctg caa gtg ctg cgc gac agt aac gac ggt ttt gtg att gcg caa 1872
Arg Leu Gln Val Leu Arg Asp Ser Asn Asp Gly Phe Val Ile Ala Gln
610 615 620

aaa gat ctg gag att cgc ggc cct ggc gaa ttg tta ggc acg cgt cag 1920
Lys Asp Leu Glu Ile Arg Gly Pro Gly Glu Leu Leu Gly Thr Arg Gln
625 630 635 640

acg ggt aat gct gaa ttt aaa gtg gcg gat tta ctg cgc gat cag gcg 1968
Thr Gly Asn Ala Glu Phe Lys Val Ala Asp Leu Leu Arg Asp Gln Ala
645 650 655

atg atc ccg gaa gtt cag cgc ctg gca cgc cat att cac gaa cgt tac 2016
Met Ile Pro Glu Val Gln Arg Leu Ala Arg His Ile His Glu Arg Tyr
660 665 670

cca caa cag gca aaa gcc ctg ata gaa cgc tgg atg ccg gag acg gaa 2064
Pro Gln Gln Ala Lys Ala Leu Ile Glu Arg Trp Met Pro Glu Thr Glu
675 680 685

cgt tac tcg aat gcg taa 2082
Arg Tyr Ser Asn Ala *
690

<210> 201
<211> 1287
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1287)

<400> 201
atg aaa acc tct ctg ttt aaa agc ctt tac ttt cag gtc ctg aca gcg 48

Met Lys Thr Ser Leu Phe Lys Ser Leu Tyr Phe Gln Val Leu Thr Ala
1 5 10 15

ata gcc att ggt att ctc ctt ggc cat ttc tat cct gaa ata ggc gag 96
Ile Ala Ile Gly Ile Leu Leu Gly His Phe Tyr Pro Glu Ile Gly Glu
20 25 30

caa atg aaa ccg ctt ggc gac ggc ttc gtt aag ctc att aag atg atc 144
Gln Met Lys Pro Leu Gly Asp Gly Phe Val Lys Leu Ile Lys Met Ile
35 40 45

atc gct cct gtc atc ttt tgt acc gtc gta acg ggc att gcg ggc atg 192
Ile Ala Pro Val Ile Phe Cys Thr Val Val Thr Gly Ile Ala Gly Met
50 55 60

gaa agc atg aag gcg gtc ggt cgt acc ggc gca gtc gca ctg ctt tac 240
Glu Ser Met Lys Ala Val Gly Arg Thr Gly Ala Val Ala Leu Leu Tyr
65 70 75 80

ttt gaa att gtc agt acc atc gcg ctg att att ggt ctt atc atc gtt 288
Phe Glu Ile Val Ser Thr Ile Ala Leu Ile Ile Gly Leu Ile Ile Val
85 90 95

aac gtc gtg cag cct ggt gcc gga atg aac gtc gat ccg gca acg ctt 336
Asn Val Val Gln Pro Gly Ala Gly Met Asn Val Asp Pro Ala Thr Leu
100 105 110

gat gcg aaa gcg gta gcg gtt tac gcc gat cag gcg aaa gac cag ggc 384
Asp Ala Lys Ala Val Ala Val Tyr Ala Asp Gln Ala Lys Asp Gln Gly
115 120 125

att gtc gcc ttc att atg gat gtc atc ccg gcg agc gtc att ggc gca 432
Ile Val Ala Phe Ile Met Asp Val Ile Pro Ala Ser Val Ile Gly Ala
130 135 140

ttt gcc agc ggt aac att ctg cag gtg ctg ctg ttt gcc gta ctg ttt 480
Phe Ala Ser Gly Asn Ile Leu Gln Val Leu Leu Phe Ala Val Leu Phe
145 150 155 160

ggt ttt gcg ctc cac cgt ctg ggc agc aaa ggc caa ctg att ttt aac 528
Gly Phe Ala Leu His Arg Leu Gly Ser Lys Gly Gln Leu Ile Phe Asn
165 170 175

gtc atc gaa agt ttc tcg cag gtc atc ttc ggc atc atc aat atg atc 576
Val Ile Glu Ser Phe Ser Gln Val Ile Phe Gly Ile Ile Asn Met Ile
180 185 190

atg cgt ctg gca cct att ggt gcg ttc ggg gca atg gcg ttt acc atc 624
Met Arg Leu Ala Pro Ile Gly Ala Phe Gly Ala Met Ala Phe Thr Ile
195 200 205

ggt aaa tac ggc gtc ggc aca ctg gtg caa ctg ggg cag ctg att atc 672
Gly Lys Tyr Gly Val Gly Thr Leu Val Gln Leu Gly Gln Leu Ile Ile
210 215 220

tgt ttc tac att acc tgt atc ctg ttt gtg gtg ctg gta ttg ggt tca 720
Cys Phe Tyr Ile Thr Cys Ile Leu Phe Val Val Leu Val Leu Gly Ser
225 230 235 240

atc gct aaa gcg act ggt ttc agt atc ttc aaa ttt atc cgc tac atc 768
Ile Ala Lys Ala Thr Gly Phe Ser Ile Phe Lys Phe Ile Arg Tyr Ile

245 250 255

cgt gaa gaa ctg ctg att gta ctg ggg act tca tct tcc gag tcg gcg 816
Arg Glu Glu Leu Leu Ile Val Leu Gly Thr Ser Ser Ser Glu Ser Ala
260 265 270

ctg ccg cgt atg ctc gac aag atg gag aaa ctc ggc tgc cgt aaa tcg 864
Leu Pro Arg Met Leu Asp Lys Met Glu Lys Leu Gly Cys Arg Lys Ser
275 280 285

gtg gtg ggg ctg gtc atc ccg aca ggc tac tcg ttt aac ctt gat ggc 912
Val Val Gly Leu Val Ile Pro Thr Gly Tyr Ser Phe Asn Leu Asp Gly
290 295 300

aca tcg ata tac ctg aca atg gcg gcg gtg ttt atc gcc cag gcc act 960
Thr Ser Ile Tyr Leu Thr Met Ala Ala Val Phe Ile Ala Gln Ala Thr
305 310 315 320

aac agt cag atg gat atc gtc cac caa atc acg ctg tta atc gtg ttg 1008
Asn Ser Gln Met Asp Ile Val His Gln Ile Thr Leu Leu Ile Val Leu
325 330 335

ctg ctt tct tct aaa ggg gcg gca ggg gta acg ggt agt ggc ttt atc 1056
Leu Leu Ser Ser Lys Gly Ala Ala Gly Val Thr Gly Ser Gly Phe Ile
340 345 350

gtg ctg gcg gcg acg ctc tct gcg gtg ggc cat ttg ccg gta gcg ggt 1104
Val Leu Ala Ala Thr Leu Ser Ala Val Gly His Leu Pro Val Ala Gly
355 360 365

ctg gcg ctg atc ctc ggt atc gac cgc ttt atg tca gaa gct cgt gcg 1152
Leu Ala Leu Ile Leu Gly Ile Asp Arg Phe Met Ser Glu Ala Arg Ala
370 375 380

ctg act aac ctg gtc ggt aac ggc gta gcg acc att gtc gtt gct aag 1200
Leu Thr Asn Leu Val Gly Asn Gly Val Ala Thr Ile Val Val Ala Lys
385 390 395 400

tgg gtg aaa gaa ctg gac cac aaa aaa ctg gac gat gtg ctg aat aat 1248
Trp Val Lys Glu Leu Asp His Lys Lys Leu Asp Asp Val Leu Asn Asn
405 410 415

cgt gcg ccg gat ggc aaa acg cac gaa tta tcc tct taa 1287
Arg Ala Pro Asp Gly Lys Thr His Glu Leu Ser Ser *
420 425

<210> 202
<211> 1191
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1191)

<400> 202
gtg acc acc cga cag cat tcg tcg ttt gct att gtt ttt atc ctt ggc 48
Met Thr Thr Arg Gln His Ser Ser Phe Ala Ile Val Phe Ile Leu Gly
1 5 10 15

ctg ctg gcc atg ttg atg ccg ctg tcg att gat atg tat ctg ccc gcg 96
Leu Leu Ala Met Leu Met Pro Leu Ser Ile Asp Met Tyr Leu Pro Ala
20 25 30

cta ccg gta att tca gcg cag ttt ggc gta ccg gcg ggc agt acg cag 144
Leu Pro Val Ile Ser Ala Gln Phe Gly Val Pro Ala Gly Ser Thr Gln
35 40 45

atg acc ctc agt act tat att ctg ggc ttt gcg ttg ggg cag tta atc 192
Met Thr Leu Ser Thr Tyr Ile Leu Gly Phe Ala Leu Gly Gln Leu Ile
50 55 60

tac ggg ccg atg gca gac agc ttc ggg cgt aag ccg gtg gtg ctc ggc 240
Tyr Gly Pro Met Ala Asp Ser Phe Gly Arg Lys Pro Val Val Leu Gly
65 70 75 80

ggt acg ctg gtg ttt gcc gcc gcc gcg gtg gcg tgt gcg ttg gca aac 288
Gly Thr Leu Val Phe Ala Ala Ala Ala Val Ala Cys Ala Leu Ala Asn
85 90 95

acc atc gat cag ctg att gtg atg cgt ttc ttc cac ggg ctg gct gcg 336
Thr Ile Asp Gln Leu Ile Val Met Arg Phe Phe His Gly Leu Ala Ala
100 105 110

gct gcg gcc agc gtg gtc att aac gcc ctg atg cgc gat att tac ccg 384
Ala Ala Ala Ser Val Val Ile Asn Ala Leu Met Arg Asp Ile Tyr Pro
115 120 125

aaa gaa gag ttc tcg cgg atg atg tcg ttt gtc atg ctg gtg aca acc 432
Lys Glu Glu Phe Ser Arg Met Met Ser Phe Val Met Leu Val Thr Thr
130 135 140

att gca ccg ctg atg gca ccg ata gtt ggc ggc tgg gtg ctg gtg tgg 480
Ile Ala Pro Leu Met Ala Pro Ile Val Gly Gly Trp Val Leu Val Trp
145 150 155 160

ctg agc tgg cat tac atc ttc tgg atc ctg gca tta gcg gcg att ctg 528
Leu Ser Trp His Tyr Ile Phe Trp Ile Leu Ala Leu Ala Ala Ile Leu
165 170 175

gct tcg gca atg att ttc ttc ctg att aaa gaa acc tta cca ccg gag 576
Ala Ser Ala Met Ile Phe Phe Leu Ile Lys Glu Thr Leu Pro Pro Glu
180 185 190

cgt cgt cag cca ttt cac att cgt acc act att ggt aac ttt gcg gcg 624
Arg Arg Gln Pro Phe His Ile Arg Thr Thr Ile Gly Asn Phe Ala Ala
195 200 205

ctg ttc cgc cat aaa cgt gtc ctg agc tac atg ctt gcc agt ggt ttc 672
Leu Phe Arg His Lys Arg Val Leu Ser Tyr Met Leu Ala Ser Gly Phe
210 215 220

agc ttt gcc ggg atg ttc tca ttc tta agc gcc gga ccg ttt gtt tat 720
Ser Phe Ala Gly Met Phe Ser Phe Leu Ser Ala Gly Pro Phe Val Tyr
225 230 235 240

att gaa att aac cac gtc gcg ccg gaa aac ttt ggt tat tac ttt gcg 768
Ile Glu Ile Asn His Val Ala Pro Glu Asn Phe Gly Tyr Tyr Phe Ala
245 250 255

cta aac att gtt ttt ctg ttc gtg atg acc atc ttt aac agc cgc ttc 816

Leu Asn Ile Val Phe Leu Phe Val Met Thr Ile Phe Asn Ser Arg Phe
260 265 270

gtc cgc cgc att ggc gcg tta aat atg ttc cgc tcg ggg ttg tgg ata 864
Val Arg Arg Ile Gly Ala Leu Asn Met Phe Arg Ser Gly Leu Trp Ile
275 280 285

caa ttt att atg gca gcg tgg atg gtc atc agt gcg ctg ctg ggg ctg 912
Gln Phe Ile Met Ala Ala Trp Met Val Ile Ser Ala Leu Leu Gly Leu
290 295 300

gga ttt tgg tcg ctg gtg gtt ggc gtt gcg gcg ttt gtg ggc tgc gtg 960
Gly Phe Trp Ser Leu Val Val Gly Val Ala Ala Phe Val Gly Cys Val
305 310 315 320

tcg atg gtg tca tcc aat gcg atg gcg gtc att ctt gat gag ttt ccc 1008
Ser Met Val Ser Ser Asn Ala Met Ala Val Ile Leu Asp Glu Phe Pro
325 330 335

cat atg gcg gga acg gca tct tcg ctg gca gga acc ttc cgt ttt ggc 1056
His Met Ala Gly Thr Ala Ser Ser Leu Ala Gly Thr Phe Arg Phe Gly
340 345 350

ata ggg gca att gtt ggc gca ttg ctt tct ctt gcg acc ttt aac tct 1104
Ile Gly Ala Ile Val Gly Ala Leu Leu Ser Leu Ala Thr Phe Asn Ser
355 360 365

gca tgg ccg atg att tgg tca att gca ttc tgc gca acc agc tcc att 1152
Ala Trp Pro Met Ile Trp Ser Ile Ala Phe Cys Ala Thr Ser Ser Ile
370 375 380

ctc ttc tgt ctg tac gcc agt cgg ccg aaa aaa cgg tga 1191
Leu Phe Cys Leu Tyr Ala Ser Arg Pro Lys Lys Arg *
385 390 395

<210> 203
<211> 696
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(696)

<400> 203
atg cga ctt gat aaa ttt atc gca cag caa ctc ggc gtt agc cgt gct 48
Met Arg Leu Asp Lys Phe Ile Ala Gln Gln Leu Gly Val Ser Arg Ala
1 5 10 15

att gcc ggg cgt gaa atc cgc ggc aat cgt gtc acc gtc gat ggc gaa 96
Ile Ala Gly Arg Glu Ile Arg Gly Asn Arg Val Thr Val Asp Gly Glu
20 25 30

atc gtc cgt aat gca gcg ttc aaa ctg ctt cct gaa cat gat gtc gct 144
Ile Val Arg Asn Ala Ala Phe Lys Leu Leu Pro Glu His Asp Val Ala
35 40 45

tac gat ggc aac ccg ctg gcg cag caa cac ggt cca cgt tac ttc atg 192
Tyr Asp Gly Asn Pro Leu Ala Gln Gln His Gly Pro Arg Tyr Phe Met
50 55 60

```
ctc aat aag cct cag ggc tat gtt tgc tcc acg gac gac cct gat cac    240
Leu Asn Lys Pro Gln Gly Tyr Val Cys Ser Thr Asp Asp Pro Asp His
 65                 70                  75                  80

cca acg gtg ctc tat ttt ctt gat gaa ccg gta gcg tgg aaa ctg cat    288
Pro Thr Val Leu Tyr Phe Leu Asp Glu Pro Val Ala Trp Lys Leu His
                85                  90                  95

gcg gcg ggg cgg ttg gat att gat acc acc ggt ctg gtg ctg atg act    336
Ala Ala Gly Arg Leu Asp Ile Asp Thr Thr Gly Leu Val Leu Met Thr
            100                 105                 110

gat gat ggt cag tgg tcg cac cgc att act tct ccg cgc cat cat tgc    384
Asp Asp Gly Gln Trp Ser His Arg Ile Thr Ser Pro Arg His His Cys
        115                 120                 125

gag aag acc tat ctg gtg aca ctg gaa tca cct gta gct gac gat acg    432
Glu Lys Thr Tyr Leu Val Thr Leu Glu Ser Pro Val Ala Asp Asp Thr
    130                 135                 140

gca gag caa ttt gct aaa ggc gtg cag ctg cat aac gaa aaa gat ctc    480
Ala Glu Gln Phe Ala Lys Gly Val Gln Leu His Asn Glu Lys Asp Leu
145                 150                 155                 160

act aag cct gcg gtg ctg gaa gtg att acc cca acg cag gtt cgt ctg    528
Thr Lys Pro Ala Val Leu Glu Val Ile Thr Pro Thr Gln Val Arg Leu
                165                 170                 175

acc atc agc gaa ggg cgt tat cat cag gtg aaa cgc atg ttc gcc gcc    576
Thr Ile Ser Glu Gly Arg Tyr His Gln Val Lys Arg Met Phe Ala Ala
            180                 185                 190

gtg ggt aac cac gtg gtt gag ctg cat cgt gaa cgt att ggc ggt att    624
Val Gly Asn His Val Val Glu Leu His Arg Glu Arg Ile Gly Gly Ile
        195                 200                 205

acg ctg gat gct gat tta gcc ccc ggt gaa tat cgt ccg tta act gaa    672
Thr Leu Asp Ala Asp Leu Ala Pro Gly Glu Tyr Arg Pro Leu Thr Glu
    210                 215                 220

gaa gaa att gcc agc gtc gtc taa                                    696
Glu Glu Ile Ala Ser Val Val  *
225                 230
```

<210> 204
<211> 1329
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1329)

<400> 204

```
atg aaa aaa att gaa tgc gct tgc aat ttt ctg atg gat aaa gat gcg    48
Met Lys Lys Ile Glu Cys Ala Cys Asn Phe Leu Met Asp Lys Asp Ala
 1               5                  10                  15

cag ggg tat atc gac ctg tct gat ttg gat tta aca agt tgt cat ttt    96
```

```
Gln Gly Tyr Ile Asp Leu Ser Asp Leu Asp Leu Thr Ser Cys His Phe
            20                  25                  30

aaa ggt gac gtt ata tcg aag gtg tct ttt tta tca tca aat cta caa      144
Lys Gly Asp Val Ile Ser Lys Val Ser Phe Leu Ser Ser Asn Leu Gln
        35                  40                  45

cat gta aca ttc gaa tgt aaa gaa att ggg gat tgc aat ttt act act      192
His Val Thr Phe Glu Cys Lys Glu Ile Gly Asp Cys Asn Phe Thr Thr
    50                  55                  60

gca ata gtt gat aat gtc ata ttt aga tgt cga cgt tta cac aat gtg      240
Ala Ile Val Asp Asn Val Ile Phe Arg Cys Arg Arg Leu His Asn Val
65                  70                  75                  80

att ttt atc aaa gcg agt ggt gaa tgt gtc gat ttc agc aaa aat att      288
Ile Phe Ile Lys Ala Ser Gly Glu Cys Val Asp Phe Ser Lys Asn Ile
                85                  90                  95

ctt gat aca gtt gac ttc tcg cag agt caa ctt ggt cat agt aat ttt      336
Leu Asp Thr Val Asp Phe Ser Gln Ser Gln Leu Gly His Ser Asn Phe
            100                 105                 110

cgc gaa tgt cag att aga aat tca aac ttc gat aat tgt tat ctt tac      384
Arg Glu Cys Gln Ile Arg Asn Ser Asn Phe Asp Asn Cys Tyr Leu Tyr
        115                 120                 125

gct tcg cac ttc acc aga gca gag ttt ctg tct gcc aaa gaa ata tca      432
Ala Ser His Phe Thr Arg Ala Glu Phe Leu Ser Ala Lys Glu Ile Ser
    130                 135                 140

ttt att aaa tcg aat ttg aca gct gtt atg ttt gat tat gtg cga atg      480
Phe Ile Lys Ser Asn Leu Thr Ala Val Met Phe Asp Tyr Val Arg Met
145                 150                 155                 160

tcg aca ggg aat ttt aaa gat tgc att aca gaa caa ttg gaa tta act      528
Ser Thr Gly Asn Phe Lys Asp Cys Ile Thr Glu Gln Leu Glu Leu Thr
                165                 170                 175

att gat tat tca gat ata ttt tgg aat gaa gat ctc gat ggt tat atc      576
Ile Asp Tyr Ser Asp Ile Phe Trp Asn Glu Asp Leu Asp Gly Tyr Ile
            180                 185                 190

aat aac att ata aaa atg att gat aca ttg cca gat aat gca atg ata      624
Asn Asn Ile Ile Lys Met Ile Asp Thr Leu Pro Asp Asn Ala Met Ile
        195                 200                 205

ttg aaa tcc gtt ctg gcc gta aaa ctg gtg atg caa tta aaa ata ctt      672
Leu Lys Ser Val Leu Ala Val Lys Leu Val Met Gln Leu Lys Ile Leu
    210                 215                 220

aat att gtt aat aaa aac ttt att gag aat atg aag aaa ata ttt agc      720
Asn Ile Val Asn Lys Asn Phe Ile Glu Asn Met Lys Lys Ile Phe Ser
225                 230                 235                 240

cat tgt cct tat ata aaa gat ccc att ata cgc agt tat atc cat tct      768
His Cys Pro Tyr Ile Lys Asp Pro Ile Ile Arg Ser Tyr Ile His Ser
                245                 250                 255

gat gaa gat aac aag ttc gat gat ttt atg cgt caa cat cga ttc agt      816
Asp Glu Asp Asn Lys Phe Asp Asp Phe Met Arg Gln His Arg Phe Ser
```

```
            260                 265                 270

gag gtg aat ttc gat acc caa cag atg atc gat ttt att aac aga ttt     864
Glu Val Asn Phe Asp Thr Gln Gln Met Ile Asp Phe Ile Asn Arg Phe
        275                 280                 285

aat acg aat aaa tgg cta att gat aaa aat aac aat ttt ttt atc caa     912
Asn Thr Asn Lys Trp Leu Ile Asp Lys Asn Asn Asn Phe Phe Ile Gln
    290                 295                 300

ctt atc gat cag gcc tta cga tca acg gat gat atg atc aaa gca aat     960
Leu Ile Asp Gln Ala Leu Arg Ser Thr Asp Asp Met Ile Lys Ala Asn
305                 310                 315                 320

gtt tgg cat ctt tat aaa gag tgg att cgt agt gat gat gtt tca cct     1008
Val Trp His Leu Tyr Lys Glu Trp Ile Arg Ser Asp Asp Val Ser Pro
                325                 330                 335

ata ttt ata gaa act gaa gat aat tta aga acc ttt aac acg aat gaa     1056
Ile Phe Ile Glu Thr Glu Asp Asn Leu Arg Thr Phe Asn Thr Asn Glu
            340                 345                 350

tta aca cga aac gat aat atc ttt atc ctg ttc tcc tca gtc gat gat     1104
Leu Thr Arg Asn Asp Asn Ile Phe Ile Leu Phe Ser Ser Val Asp Asp
        355                 360                 365

ggg cca gtt atg gtg gta agc tcc cag cgc tta cat gat atg ttg aat     1152
Gly Pro Val Met Val Val Ser Ser Gln Arg Leu His Asp Met Leu Asn
    370                 375                 380

cct aca aaa gat acc aat tgg aat tcc acg tat atc tac aaa tcc aga     1200
Pro Thr Lys Asp Thr Asn Trp Asn Ser Thr Tyr Ile Tyr Lys Ser Arg
385                 390                 395                 400

cat gag atg ttg cct gtt aat ctt act cag gaa aca ctt ttc agc tcc     1248
His Glu Met Leu Pro Val Asn Leu Thr Gln Glu Thr Leu Phe Ser Ser
                405                 410                 415

aaa tct cat ggt aaa tat gcg ctt ttc ccc att ttt act gcg agt tgg     1296
Lys Ser His Gly Lys Tyr Ala Leu Phe Pro Ile Phe Thr Ala Ser Trp
            420                 425                 430

cga gct cat cgt ata atg aat aag ggt gtt taa                         1329
Arg Ala His Arg Ile Met Asn Lys Gly Val  *
        435                 440
```

<210> 205
<211> 717
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(717)

<400> 205

```
gtg gga cgt aaa tgg gcc aat att gtt gct aaa aaa acg gct aaa gac     48
Met Gly Arg Lys Trp Ala Asn Ile Val Ala Lys Lys Thr Ala Lys Asp
 1               5                  10                  15
```

```
ggt gca acg tct aaa att tat gca aaa ttc ggt gta gaa atc tat gct    96
Gly Ala Thr Ser Lys Ile Tyr Ala Lys Phe Gly Val Glu Ile Tyr Ala
            20                  25                  30

gct gct aaa caa ggt gaa ccc gat cca gaa tta aac aca tct tta aaa   144
Ala Ala Lys Gln Gly Glu Pro Asp Pro Glu Leu Asn Thr Ser Leu Lys
        35                  40                  45

ttc gtt att gaa cgt gca aag cag gca caa gtt cca aag cac gtt att   192
Phe Val Ile Glu Arg Ala Lys Gln Ala Gln Val Pro Lys His Val Ile
    50                  55                  60

gat aaa gca att gat aaa gcc aaa ggc ggc gga gat gaa acg ttc gtg   240
Asp Lys Ala Ile Asp Lys Ala Lys Gly Gly Gly Asp Glu Thr Phe Val
65                  70                  75                  80

cag gga cgt tat gaa ggc ttt ggt cct aat ggc tca atg att atc gcc   288
Gln Gly Arg Tyr Glu Gly Phe Gly Pro Asn Gly Ser Met Ile Ile Ala
                85                  90                  95

gag aca ttg act tca aat gtt aac cgt acg att gct aac gtt cgc aca   336
Glu Thr Leu Thr Ser Asn Val Asn Arg Thr Ile Ala Asn Val Arg Thr
            100                 105                 110

att ttc aat aaa aaa ggc ggc aat atc gga gcg gca ggt tct gtc agc   384
Ile Phe Asn Lys Lys Gly Gly Asn Ile Gly Ala Ala Gly Ser Val Ser
        115                 120                 125

tat atg ttt gac aat acg ggt gtg att gta ttt aaa ggg aca gac cct   432
Tyr Met Phe Asp Asn Thr Gly Val Ile Val Phe Lys Gly Thr Asp Pro
    130                 135                 140

gac cat att ttt gaa att tta ctt gaa gct gaa gtt gat gtt cgt gat   480
Asp His Ile Phe Glu Ile Leu Leu Glu Ala Glu Val Asp Val Arg Asp
145                 150                 155                 160

gtg act gaa gaa gaa ggt aac att gtt att tat act gaa cct act gac   528
Val Thr Glu Glu Glu Gly Asn Ile Val Ile Tyr Thr Glu Pro Thr Asp
                165                 170                 175

ctt cat aaa gga atc gcg gct cta aaa gca gct gga atc act gag ttc   576
Leu His Lys Gly Ile Ala Ala Leu Lys Ala Ala Gly Ile Thr Glu Phe
            180                 185                 190

tca aca aca gaa tta gaa atg att gct caa tct gaa gtt gag ctt tcc   624
Ser Thr Thr Glu Leu Glu Met Ile Ala Gln Ser Glu Val Glu Leu Ser
        195                 200                 205

cca gaa gat tta gaa atc ttt gaa ggg ctt gtt gat gcc ctt gaa gat   672
Pro Glu Asp Leu Glu Ile Phe Glu Gly Leu Val Asp Ala Leu Glu Asp
    210                 215                 220

gac gac gat gta caa aaa gtt tat cat aac gtc gca aat ctc taa       717
Asp Asp Asp Val Gln Lys Val Tyr His Asn Val Ala Asn Leu  *
225                 230                 235
```

<210> 206
<211> 786
<212> DNA

<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(786)

<400> 206
gtg gtt ctt atg tcc gaa aca aaa aac gaa ctt gaa gac ctg ctg gaa 48
Met Val Leu Met Ser Glu Thr Lys Asn Glu Leu Glu Asp Leu Leu Glu
1 5 10 15

aaa gca gca act gaa ccg gcg cac cgc ccg gcc ttt ttc cgt act cta 96
Lys Ala Ala Thr Glu Pro Ala His Arg Pro Ala Phe Phe Arg Thr Leu
20 25 30

ctg gaa tcc acc gtc tgg gtg cct ggt acg gcg gcg cag ggc gag gct 144
Leu Glu Ser Thr Val Trp Val Pro Gly Thr Ala Ala Gln Gly Glu Ala
35 40 45

gtg gtt gaa gat agc gcg ctt gat tta cag cac tgg gaa aaa gaa gac 192
Val Val Glu Asp Ser Ala Leu Asp Leu Gln His Trp Glu Lys Glu Asp
50 55 60

ggc acc agc gtc att cct ttt ttc acc tcg tta gaa gca ctt caa cag 240
Gly Thr Ser Val Ile Pro Phe Phe Thr Ser Leu Glu Ala Leu Gln Gln
65 70 75 80

gcg gtt gaa gac gaa cag gca ttt gtc gta atg ccc gtt cgc acg ctg 288
Ala Val Glu Asp Glu Gln Ala Phe Val Val Met Pro Val Arg Thr Leu
85 90 95

ttt gag atg aca ctt ggc gaa acg ctc ttc ctt aat gcc aaa ctg cca 336
Phe Glu Met Thr Leu Gly Glu Thr Leu Phe Leu Asn Ala Lys Leu Pro
100 105 110

acc ggt aaa gaa ttt atg ccg cgt gaa atc agt ttg ttg att ggt gaa 384
Thr Gly Lys Glu Phe Met Pro Arg Glu Ile Ser Leu Leu Ile Gly Glu
115 120 125

gag gga aat ccg ctg agc agc cag gaa atc ctg gaa ggc ggt gaa tcg 432
Glu Gly Asn Pro Leu Ser Ser Gln Glu Ile Leu Glu Gly Gly Glu Ser
130 135 140

ctg ata tta tcg gaa gtc gca gag ccg cca gca caa atg att gat tca 480
Leu Ile Leu Ser Glu Val Ala Glu Pro Pro Ala Gln Met Ile Asp Ser
145 150 155 160

ctc acc acc tta ttt aaa acc att aag ccg gtg aag cgt gct ttt att 528
Leu Thr Thr Leu Phe Lys Thr Ile Lys Pro Val Lys Arg Ala Phe Ile
165 170 175

tgt tca att aaa gag aac gaa gag gca cag cct aat tta ctt att ggc 576
Cys Ser Ile Lys Glu Asn Glu Glu Ala Gln Pro Asn Leu Leu Ile Gly
180 185 190

att gaa gcc gat ggt gat atc gaa gaa att att cag gcg acg gga agt 624
Ile Glu Ala Asp Gly Asp Ile Glu Glu Ile Ile Gln Ala Thr Gly Ser
195 200 205

gta gcg acc gat aca tta cct ggc gat gaa cca atc gat att tgt cag 672
Val Ala Thr Asp Thr Leu Pro Gly Asp Glu Pro Ile Asp Ile Cys Gln

210 215 220

gtg aaa aaa ggg gaa aaa gga att agc cac ttt att acc gaa cat att 720
Val Lys Lys Gly Glu Lys Gly Ile Ser His Phe Ile Thr Glu His Ile
225 230 235 240

gcg cca ttc tat gaa cgt cgc tgg ggt ggt ttt ttg cgt gac ttt aaa 768
Ala Pro Phe Tyr Glu Arg Arg Trp Gly Gly Phe Leu Arg Asp Phe Lys
245 250 255

cag aat cgg ata atc taa 786
Gln Asn Arg Ile Ile *
260

<210> 207
<211> 1266
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1266)

<400> 207
atg ctc acg aaa aag aaa tgg gcg tta ttt agt cta tta aca ctg tgt 48
Met Leu Thr Lys Lys Lys Trp Ala Leu Phe Ser Leu Leu Thr Leu Cys
1 5 10 15

ggc ggt aca att tat aaa tta ccg tcg ctg aaa gat gcg ttt tat atc 96
Gly Gly Thr Ile Tyr Lys Leu Pro Ser Leu Lys Asp Ala Phe Tyr Ile
20 25 30

ccg atg cag gaa tat ttc cat ttg acc aat ggt caa att ggt aat gct 144
Pro Met Gln Glu Tyr Phe His Leu Thr Asn Gly Gln Ile Gly Asn Ala
35 40 45

atg tcg gta aac tca ttt gtc acc aca gtg ggc ttt ttt ctg tct att 192
Met Ser Val Asn Ser Phe Val Thr Thr Val Gly Phe Phe Leu Ser Ile
50 55 60

tat ttt gcc gat aaa cta ccg cgc aga tac acc atg tca ttc tca ctc 240
Tyr Phe Ala Asp Lys Leu Pro Arg Arg Tyr Thr Met Ser Phe Ser Leu
65 70 75 80

att gcg aca gga tta ctg ggt gtt tat ttg acg aca atg ccg ggg tat 288
Ile Ala Thr Gly Leu Leu Gly Val Tyr Leu Thr Thr Met Pro Gly Tyr
85 90 95

tgg ggc atc ctc ttt gtc tgg gcg cta ttt ggc gtt act tgc gac atg 336
Trp Gly Ile Leu Phe Val Trp Ala Leu Phe Gly Val Thr Cys Asp Met
100 105 110

atg aac tgg ccg gtc ttg ctc aag tcg gta agt cga ttg ggc aat agc 384
Met Asn Trp Pro Val Leu Leu Lys Ser Val Ser Arg Leu Gly Asn Ser
115 120 125

gaa caa caa ggt cgg ttg ttt ggc ttc ttc gaa aca ggg cgt ggc att 432
Glu Gln Gln Gly Arg Leu Phe Gly Phe Phe Glu Thr Gly Arg Gly Ile
130 135 140

gtc gat acc gtg gtg gca ttt tct gcg ttg gca gta ttt acc tgg ttt 480
Val Asp Thr Val Val Ala Phe Ser Ala Leu Ala Val Phe Thr Trp Phe
145 150 155 160

ggc agt ggc tta tta ggt ttt aaa gca ggc atc tgg ttc tat tcc ctt 528
Gly Ser Gly Leu Leu Gly Phe Lys Ala Gly Ile Trp Phe Tyr Ser Leu
165 170 175

att gtg att gcc gta ggc att att att ttc ttt gtc ctg aat gac aaa 576
Ile Val Ile Ala Val Gly Ile Ile Ile Phe Phe Val Leu Asn Asp Lys
180 185 190

gaa gag gca ccg tcc gtt gag gtg aaa aaa gaa gac gga gca tcg aaa 624
Glu Glu Ala Pro Ser Val Glu Val Lys Lys Glu Asp Gly Ala Ser Lys
195 200 205

aac acc agt atg acc tcg gtg ctg aaa gac aaa act atc tgg ctt atc 672
Asn Thr Ser Met Thr Ser Val Leu Lys Asp Lys Thr Ile Trp Leu Ile
210 215 220

gct ttt aac gtc ttc ttc gtt tac gcg gtt tac tgt ggc ctg aca ttc 720
Ala Phe Asn Val Phe Phe Val Tyr Ala Val Tyr Cys Gly Leu Thr Phe
225 230 235 240

ttc att cca ttc ctg aaa aac atc tat cta ttg ccc gtt gcg ctg gtg 768
Phe Ile Pro Phe Leu Lys Asn Ile Tyr Leu Leu Pro Val Ala Leu Val
245 250 255

ggg gct tac ggc atc att aac caa tac tgt ctg aaa atg att ggt gga 816
Gly Ala Tyr Gly Ile Ile Asn Gln Tyr Cys Leu Lys Met Ile Gly Gly
260 265 270

ccg att ggt ggc atg att tca gat aag atc ctg aaa tcg cca agt aaa 864
Pro Ile Gly Gly Met Ile Ser Asp Lys Ile Leu Lys Ser Pro Ser Lys
275 280 285

tat cta tgc tac acc ttt atc att agt acc gct gcg ctc gta ctg ttg 912
Tyr Leu Cys Tyr Thr Phe Ile Ile Ser Thr Ala Ala Leu Val Leu Leu
290 295 300

att atg ctg ccg cac gaa agt atg ccg gtc tat tta ggg atg gca tgt 960
Ile Met Leu Pro His Glu Ser Met Pro Val Tyr Leu Gly Met Ala Cys
305 310 315 320

acg ctg ggc ttt ggc gcg ata gtc ttt aca cag cga gcc gta ttt ttt 1008
Thr Leu Gly Phe Gly Ala Ile Val Phe Thr Gln Arg Ala Val Phe Phe
325 330 335

gca cct atc ggc gaa gca aaa att gct gaa aat aaa aca ggt gcg gcg 1056
Ala Pro Ile Gly Glu Ala Lys Ile Ala Glu Asn Lys Thr Gly Ala Ala
340 345 350

atg gcg ttg ggt agc ttt att ggt tac gct ccg gcg atg ttc tgc ttc 1104
Met Ala Leu Gly Ser Phe Ile Gly Tyr Ala Pro Ala Met Phe Cys Phe
355 360 365

agt ctg tat ggc tac att ctg gat tta aat ccg ggg att att ggc tac 1152
Ser Leu Tyr Gly Tyr Ile Leu Asp Leu Asn Pro Gly Ile Ile Gly Tyr
370 375 380

aaa atc gtg ttt ggc att atg gcc tgc ttc gca ttc agt ggt gcg gtg 1200

```
Lys Ile Val Phe Gly Ile Met Ala Cys Phe Ala Phe Ser Gly Ala Val
385                 390                 395                 400

gtt tcc gta atg ctg gtt aag cgt att agc caa cgt aag aaa gag atg     1248
Val Ser Val Met Leu Val Lys Arg Ile Ser Gln Arg Lys Lys Glu Met
                405                 410                 415

ctg gcg gct gaa gct taa                                             1266
Leu Ala Ala Glu Ala  *
            420
```

<210> 208
<211> 2313
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(2313)

<400> 208

```
atg cct cgc ttg tta acc aaa cgc ggc tgc tgg ata acg ttg gca gcc     48
Met Pro Arg Leu Leu Thr Lys Arg Gly Cys Trp Ile Thr Leu Ala Ala
 1               5                  10                  15

gcg ccc ttt ctc ctt ttt ctt gca gcg tgg gga gca gat aaa ctc tgg     96
Ala Pro Phe Leu Leu Phe Leu Ala Ala Trp Gly Ala Asp Lys Leu Trp
            20                  25                  30

cct cta ccg ctg cat gaa gtc aat ccc gca cga gtg gtc gtg gcg cag     144
Pro Leu Pro Leu His Glu Val Asn Pro Ala Arg Val Val Val Ala Gln
        35                  40                  45

gat ggt acg ccg ctc tgg cgc ttc gcc gat gct gac ggc atc tgg cgt     192
Asp Gly Thr Pro Leu Trp Arg Phe Ala Asp Ala Asp Gly Ile Trp Arg
    50                  55                  60

tat ccg gta aca atc gaa gat gtt tct cca cgt tac ctt gaa gcg ctg     240
Tyr Pro Val Thr Ile Glu Asp Val Ser Pro Arg Tyr Leu Glu Ala Leu
65                  70                  75                  80

atc aat tat gaa gat cgc tgg ttc tgg aag cat ccg ggg gtg aat cca     288
Ile Asn Tyr Glu Asp Arg Trp Phe Trp Lys His Pro Gly Val Asn Pro
                85                  90                  95

ttc tcg gtg gcg cgc gca gca tgg caa gat ctc act tcg gga cgg gtt     336
Phe Ser Val Ala Arg Ala Ala Trp Gln Asp Leu Thr Ser Gly Arg Val
            100                 105                 110

att tcc ggt ggc agc acg ctc act atg cag gtt gct cgt ctg ctt gat     384
Ile Ser Gly Gly Ser Thr Leu Thr Met Gln Val Ala Arg Leu Leu Asp
        115                 120                 125

cct cac ccc aaa aca ttt ggc ggc aaa att cgc cag ctc tgg cgc gcg     432
Pro His Pro Lys Thr Phe Gly Gly Lys Ile Arg Gln Leu Trp Arg Ala
    130                 135                 140

ttg caa ctg gaa tgg cat ctg tct aag cgt gaa att ctg acc ttg tat     480
Leu Gln Leu Glu Trp His Leu Ser Lys Arg Glu Ile Leu Thr Leu Tyr
145                 150                 155                 160
```

```
ctt aac cgc gct ccg ttt ggc ggt acg ttg cag ggg atc ggt gcg gca    528
Leu Asn Arg Ala Pro Phe Gly Gly Thr Leu Gln Gly Ile Gly Ala Ala
                165                 170                 175

agt tgg gct tat ctc gga aaa tcg cct gcg aat tta agc tat tcc gag    576
Ser Trp Ala Tyr Leu Gly Lys Ser Pro Ala Asn Leu Ser Tyr Ser Glu
            180                 185                 190

gcg gca atg ctg gcg gtt ttg ccg caa gcg ccc agc cgt ctt cgc ccg    624
Ala Ala Met Leu Ala Val Leu Pro Gln Ala Pro Ser Arg Leu Arg Pro
        195                 200                 205

gat cgt tgg ccg gag cgt gcc gaa gcc gcg cgt aat aaa gtg ctc gaa    672
Asp Arg Trp Pro Glu Arg Ala Glu Ala Ala Arg Asn Lys Val Leu Glu
    210                 215                 220

cgg atg gcc gtg caa ggt gtg tgg tcc cgt gag cag gta aaa gag tca    720
Arg Met Ala Val Gln Gly Val Trp Ser Arg Glu Gln Val Lys Glu Ser
225                 230                 235                 240

agg gaa gaa ccc atc tgg ctg gcc ccc cga caa atg ccg caa ctg gca    768
Arg Glu Glu Pro Ile Trp Leu Ala Pro Arg Gln Met Pro Gln Leu Ala
                245                 250                 255

ccg ctg ttt tcg cgc atg atg ctc ggt aaa agc aaa agc gac aaa atc    816
Pro Leu Phe Ser Arg Met Met Leu Gly Lys Ser Lys Ser Asp Lys Ile
            260                 265                 270

act act acg ttg gat gcc ggt ctt caa cga cgt ctg gaa gaa ctg gcg    864
Thr Thr Thr Leu Asp Ala Gly Leu Gln Arg Arg Leu Glu Glu Leu Ala
        275                 280                 285

caa aac tgg aaa ggg cgg ttg cca ccg cgc agc tca ctg gcg atg atc    912
Gln Asn Trp Lys Gly Arg Leu Pro Pro Arg Ser Ser Leu Ala Met Ile
    290                 295                 300

gtg gtt gat cat acc gat atg cgt gtt cgc ggc tgg gtg gga tcg gtt    960
Val Val Asp His Thr Asp Met Arg Val Arg Gly Trp Val Gly Ser Val
305                 310                 315                 320

gat ctc aac gat gat tca cgc ttt ggt cat gtt gat atg gtc aat tcg    1008
Asp Leu Asn Asp Asp Ser Arg Phe Gly His Val Asp Met Val Asn Ser
                325                 330                 335

atc cga tcg cca gga tca gtg ctc aaa ccg ttt gtt tat ggt ctg gcg    1056
Ile Arg Ser Pro Gly Ser Val Leu Lys Pro Phe Val Tyr Gly Leu Ala
            340                 345                 350

ctg gat gaa ggc ttg atc cac ccg gca tca ctg ctg caa gac gtc ccc    1104
Leu Asp Glu Gly Leu Ile His Pro Ala Ser Leu Leu Gln Asp Val Pro
        355                 360                 365

cgg cgc acc ggt gat tat cga cca ggt aac ttt gat agc ggt ttt cat    1152
Arg Arg Thr Gly Asp Tyr Arg Pro Gly Asn Phe Asp Ser Gly Phe His
    370                 375                 380

ggc ccg atc agc atg agc gag gcg ctg gtg cgc tcg ctg aac tta cct    1200
Gly Pro Ile Ser Met Ser Glu Ala Leu Val Arg Ser Leu Asn Leu Pro
385                 390                 395                 400
```

gct gtg cag gtg ctg gaa gcc tat gga ccg aaa cgg ttt gcg gca aag 1248
Ala Val Gln Val Leu Glu Ala Tyr Gly Pro Lys Arg Phe Ala Ala Lys
405 410 415

tta cgc aat gtt gga ttg ccg tta tat ttg ccc aac ggt gct gcg ccg 1296
Leu Arg Asn Val Gly Leu Pro Leu Tyr Leu Pro Asn Gly Ala Ala Pro
420 425 430

aat ctt tca ctc att ctc ggc ggc gct ggt gca aaa ctg gaa gat atg 1344
Asn Leu Ser Leu Ile Leu Gly Gly Ala Gly Ala Lys Leu Glu Asp Met
435 440 445

gcg gca gcg tat acc gcg ttt gct cgc cac ggc aag gca ggc aaa ttg 1392
Ala Ala Ala Tyr Thr Ala Phe Ala Arg His Gly Lys Ala Gly Lys Leu
450 455 460

cgc tta cag cct gac gat ccg ctg ctt gaa cgg cct tta atg tcg tcg 1440
Arg Leu Gln Pro Asp Asp Pro Leu Leu Glu Arg Pro Leu Met Ser Ser
465 470 475 480

ggg gcg gcg tgg atc att cgc cgg att atg gct gat gaa gcg caa ccc 1488
Gly Ala Ala Trp Ile Ile Arg Arg Ile Met Ala Asp Glu Ala Gln Pro
485 490 495

ttg ccg gat agt gcc ttg ccg cgc gtc gcc cca ctg gca tgg aaa acg 1536
Leu Pro Asp Ser Ala Leu Pro Arg Val Ala Pro Leu Ala Trp Lys Thr
500 505 510

ggc acc agc tat ggc tat cgt gac gcc tgg gcg att ggg gtt aac gct 1584
Gly Thr Ser Tyr Gly Tyr Arg Asp Ala Trp Ala Ile Gly Val Asn Ala
515 520 525

cgc tat gtc att ggg atc tgg act ggc aga ccg gac ggc acg ccc gtt 1632
Arg Tyr Val Ile Gly Ile Trp Thr Gly Arg Pro Asp Gly Thr Pro Val
530 535 540

gtt ggt cag ttt ggc ttt gcc agt gcc gta cca ttg ttg aat cag gtc 1680
Val Gly Gln Phe Gly Phe Ala Ser Ala Val Pro Leu Leu Asn Gln Val
545 550 555 560

aat aat atc tta ctg tcg cgc agt gcg aat ctg ccg gaa gac ccg cgc 1728
Asn Asn Ile Leu Leu Ser Arg Ser Ala Asn Leu Pro Glu Asp Pro Arg
565 570 575

ccg aac tca gtc act cgt ggc gtt atc tgc tgg ccg ggc gga cag tct 1776
Pro Asn Ser Val Thr Arg Gly Val Ile Cys Trp Pro Gly Gly Gln Ser
580 585 590

ttg ccg gaa ggt gac ggt aac tgt cgc cgc cgc ctg gca acc tgg ctg 1824
Leu Pro Glu Gly Asp Gly Asn Cys Arg Arg Arg Leu Ala Thr Trp Leu
595 600 605

ctg gac ggg agt cag ccg cca act cta tta ctg ccg gag cag gaa ggc 1872
Leu Asp Gly Ser Gln Pro Pro Thr Leu Leu Leu Pro Glu Gln Glu Gly
610 615 620

att aat ggc att cgt ttc ccc atc tgg ctg gat gaa aat ggc aaa cgt 1920
Ile Asn Gly Ile Arg Phe Pro Ile Trp Leu Asp Glu Asn Gly Lys Arg
625 630 635 640

gtt gcc gct gat tgc ccg caa gcg cga caa gaa atg ata aat gtc tgg 1968

Val Ala Ala Asp Cys Pro Gln Ala Arg Gln Glu Met Ile Asn Val Trp
645 650 655

ccg cta ccg ctg gaa ccc tgg ctg ccc gca tca gaa cgt cgc gct gtg 2016
Pro Leu Pro Leu Glu Pro Trp Leu Pro Ala Ser Glu Arg Arg Ala Val
660 665 670

cgc tta cca ccg gcc tcg aca agt tgc ccg cca tac ggt cac gat gcc 2064
Arg Leu Pro Pro Ala Ser Thr Ser Cys Pro Pro Tyr Gly His Asp Ala
675 680 685

cag ctc ccg ctg cag ctg aca ggc gtt cgc gat ggc gcg att att aaa 2112
Gln Leu Pro Leu Gln Leu Thr Gly Val Arg Asp Gly Ala Ile Ile Lys
690 695 700

cgt tta ccg ggc gca gcg gaa gcg act ttg ccg ttg caa tca agt gga 2160
Arg Leu Pro Gly Ala Ala Glu Ala Thr Leu Pro Leu Gln Ser Ser Gly
705 710 715 720

ggg gca ggt gaa cgc tgg tgg ttt ctg aat ggc gaa ccg tta act gaa 2208
Gly Ala Gly Glu Arg Trp Trp Phe Leu Asn Gly Glu Pro Leu Thr Glu
725 730 735

cgc ggg cgc aac gtg acc ctg cat ttg acg gat aaa ggt gat tat caa 2256
Arg Gly Arg Asn Val Thr Leu His Leu Thr Asp Lys Gly Asp Tyr Gln
740 745 750

ttg ctg gtg atg gat gac gtg ggg caa atc gcg aca gtg aaa ttt gtc 2304
Leu Leu Val Met Asp Asp Val Gly Gln Ile Ala Thr Val Lys Phe Val
755 760 765

atg caa tag 2313
Met Gln *
770

<210> 209
<211> 4962
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(4962)

<400> 209
atg aaa aag tta cgc gta gcc gcc tgc atg cta atg ctg gcg ctg gca 48
Met Lys Lys Leu Arg Val Ala Ala Cys Met Leu Met Leu Ala Leu Ala
1 5 10 15

ggg tgc gac aac aac gat aac gcg cca aca gcg gtg aaa aaa gat gcg 96
Gly Cys Asp Asn Asn Asp Asn Ala Pro Thr Ala Val Lys Lys Asp Ala
20 25 30

cct tct gaa gtt act aaa gcg gcc tct tca gaa aac gcg agt tca gca 144
Pro Ser Glu Val Thr Lys Ala Ala Ser Ser Glu Asn Ala Ser Ser Ala
35 40 45

aaa ctc tcc gtg ccg gag aga caa aaa ctg gcc caa cag agt gcc ggt 192
Lys Leu Ser Val Pro Glu Arg Gln Lys Leu Ala Gln Gln Ser Ala Gly
50 55 60

```
aag gtg ctg aca ttg ctg gat ctc tct gaa gtc caa ctt gat ggt gca    240
Lys Val Leu Thr Leu Leu Asp Leu Ser Glu Val Gln Leu Asp Gly Ala
 65                 70                  75                  80

gcc acg ctg gtg ctg acg ttc tcc atc cct ctc gac ccg gat cag gat    288
Ala Thr Leu Val Leu Thr Phe Ser Ile Pro Leu Asp Pro Asp Gln Asp
                85                  90                  95

ttc tca cgc gtt att cat gtc gtc gat aaa aaa agc ggc aaa gtg gat    336
Phe Ser Arg Val Ile His Val Val Asp Lys Lys Ser Gly Lys Val Asp
            100                 105                 110

ggt gcc tgg gag ctg tca gat aat ctt aaa gag tta cgt tta cgc cac    384
Gly Ala Trp Glu Leu Ser Asp Asn Leu Lys Glu Leu Arg Leu Arg His
        115                 120                 125

ctc gaa ccg aaa cgt gat ttg atc gtt act att ggc aag gag gtc aaa    432
Leu Glu Pro Lys Arg Asp Leu Ile Val Thr Ile Gly Lys Glu Val Lys
    130                 135                 140

gca ctc aac aac gca acc ttc agt aaa gat tac gaa aaa act ata act    480
Ala Leu Asn Asn Ala Thr Phe Ser Lys Asp Tyr Glu Lys Thr Ile Thr
145                 150                 155                 160

acc cgc gac atc caa ccc agc gtc ggt ttt gcc agc cgt ggt tcg ctg    528
Thr Arg Asp Ile Gln Pro Ser Val Gly Phe Ala Ser Arg Gly Ser Leu
                165                 170                 175

ctg cct ggc aaa gtc gtt gaa ggg ctg ccg gta atg gcg ctc aac gtt    576
Leu Pro Gly Lys Val Val Glu Gly Leu Pro Val Met Ala Leu Asn Val
            180                 185                 190

aat aat gtc gat gtt aac ttc ttc cgc gtt aag cca gaa tct ctg cca    624
Asn Asn Val Asp Val Asn Phe Phe Arg Val Lys Pro Glu Ser Leu Pro
        195                 200                 205

gca ttc att agc caa tgg gaa tac cgc aat tcg ctg gcg aac tgg cag    672
Ala Phe Ile Ser Gln Trp Glu Tyr Arg Asn Ser Leu Ala Asn Trp Gln
    210                 215                 220

tca gac aaa ctg ctg cag atg gcg gat ctg gtc tac acc gga cgg ttt    720
Ser Asp Lys Leu Leu Gln Met Ala Asp Leu Val Tyr Thr Gly Arg Phe
225                 230                 235                 240

gat ctc aat cct gcg cgt aac acc cgt gaa aaa tta ttg ctg ccg ctg    768
Asp Leu Asn Pro Ala Arg Asn Thr Arg Glu Lys Leu Leu Leu Pro Leu
                245                 250                 255

ggc gat atc aaa ccg ctt cag cag gcg ggc gtg tat ctg gct gtg atg    816
Gly Asp Ile Lys Pro Leu Gln Gln Ala Gly Val Tyr Leu Ala Val Met
            260                 265                 270

aat cag gct gga cgt tac gat tac agt aat ccc gcg acg ctg ttt acg    864
Asn Gln Ala Gly Arg Tyr Asp Tyr Ser Asn Pro Ala Thr Leu Phe Thr
        275                 280                 285

tta agt gat atc ggc gtt tca gct cac cgt tat cac aat cgt ctg gat    912
Leu Ser Asp Ile Gly Val Ser Ala His Arg Tyr His Asn Arg Leu Asp
    290                 295                 300
```

atc ttt acc caa agt ctg gaa aac ggc gcg gcc cag caa gga att gaa 960
Ile Phe Thr Gln Ser Leu Glu Asn Gly Ala Ala Gln Gln Gly Ile Glu
305 310 315 320

gtc tct tta tta aat gag aaa ggg cag act ctg act cag gca acc agt 1008
Val Ser Leu Leu Asn Glu Lys Gly Gln Thr Leu Thr Gln Ala Thr Ser
325 330 335

gac gct cag ggg cat gtg cag ctg gaa aat gat aaa aac gcg gca tta 1056
Asp Ala Gln Gly His Val Gln Leu Glu Asn Asp Lys Asn Ala Ala Leu
340 345 350

ctg ttg gcg cgt aaa gac ggt cag aca acg cta ctc gat tta aaa ctt 1104
Leu Leu Ala Arg Lys Asp Gly Gln Thr Thr Leu Leu Asp Leu Lys Leu
355 360 365

ccg gcg ctg gac tta gca gaa ttt aac att gct ggc gcg cca ggc tat 1152
Pro Ala Leu Asp Leu Ala Glu Phe Asn Ile Ala Gly Ala Pro Gly Tyr
370 375 380

agc aaa cag ttt ttc atg ttt ggc cca cgc gat ctt tat cgc cca ggt 1200
Ser Lys Gln Phe Phe Met Phe Gly Pro Arg Asp Leu Tyr Arg Pro Gly
385 390 395 400

gaa acg gta atc ctc aat ggt ttg ctg cgt gat gca gac ggt aaa gcg 1248
Glu Thr Val Ile Leu Asn Gly Leu Leu Arg Asp Ala Asp Gly Lys Ala
405 410 415

ttg ccc aat caa ccc atc aag tta gac gtg att aaa ccc gat ggg cag 1296
Leu Pro Asn Gln Pro Ile Lys Leu Asp Val Ile Lys Pro Asp Gly Gln
420 425 430

gta ctc agg agc gtc gtt agt cag ccg gag aat ggc ctc tac cac ttt 1344
Val Leu Arg Ser Val Val Ser Gln Pro Glu Asn Gly Leu Tyr His Phe
435 440 445

acc tgg cca ctc gat agc aat gcg gca acc ggt atg tgg cat att cgc 1392
Thr Trp Pro Leu Asp Ser Asn Ala Ala Thr Gly Met Trp His Ile Arg
450 455 460

gct aac acg ggc gat aat cag tat cgg atg tgg gat ttc cac gtc gaa 1440
Ala Asn Thr Gly Asp Asn Gln Tyr Arg Met Trp Asp Phe His Val Glu
465 470 475 480

gat ttt atg cca gag cgc atg gcg ctg aat ctg acc ggt gag aaa acc 1488
Asp Phe Met Pro Glu Arg Met Ala Leu Asn Leu Thr Gly Glu Lys Thr
485 490 495

ccg cta acg ccg aaa gat gaa gtg aaa ttc tcc gtg gtg ggg tac tac 1536
Pro Leu Thr Pro Lys Asp Glu Val Lys Phe Ser Val Val Gly Tyr Tyr
500 505 510

ctg tat ggt gca cct gct aat ggt aat act ttg caa ggg caa ctt ttc 1584
Leu Tyr Gly Ala Pro Ala Asn Gly Asn Thr Leu Gln Gly Gln Leu Phe
515 520 525

ctg cgc cca ctg cgt gaa gct gtg tca gcc tta cct ggt ttt gaa ttc 1632
Leu Arg Pro Leu Arg Glu Ala Val Ser Ala Leu Pro Gly Phe Glu Phe
530 535 540

ggc gat ata gct gcc gaa aat ctt tcc cgc acg ctg gat gaa gtt cag 1680

Gly Asp Ile Ala Ala Glu Asn Leu Ser Arg Thr Leu Asp Glu Val Gln
545 550 555 560

ttg acg ctg gat gat aaa ggg cgc ggc gaa gtt tct aca gaa agc cag 1728
Leu Thr Leu Asp Asp Lys Gly Arg Gly Glu Val Ser Thr Glu Ser Gln
565 570 575

tgg aag gaa acg cat tcc cca tta cag gtt att ttc cag ggt agt ttg 1776
Trp Lys Glu Thr His Ser Pro Leu Gln Val Ile Phe Gln Gly Ser Leu
580 585 590

ctg gaa tcg ggc ggt cgc ccg gtg acg cgc cgc gct gag cag gct atc 1824
Leu Glu Ser Gly Gly Arg Pro Val Thr Arg Arg Ala Glu Gln Ala Ile
595 600 605

tgg cct gcc gat gca ttg ccg ggg atc cgt ccg cag ttc gcc tcg aaa 1872
Trp Pro Ala Asp Ala Leu Pro Gly Ile Arg Pro Gln Phe Ala Ser Lys
610 615 620

tcg gtt tac gat tat cgt act gac agc acg gtg aaa cag ccc att gtt 1920
Ser Val Tyr Asp Tyr Arg Thr Asp Ser Thr Val Lys Gln Pro Ile Val
625 630 635 640

gat gaa ggc agt aac gcc gct ttt gac atc gtt tat agc gat gcg caa 1968
Asp Glu Gly Ser Asn Ala Ala Phe Asp Ile Val Tyr Ser Asp Ala Gln
645 650 655

ggc gtg aaa aaa gcc gtg tcg ggc ttg cag gtg cgc ctg att cgc gaa 2016
Gly Val Lys Lys Ala Val Ser Gly Leu Gln Val Arg Leu Ile Arg Glu
660 665 670

cgc cgc gat tac tac tgg aac tgg tca gaa gat gaa ggc tgg cag tca 2064
Arg Arg Asp Tyr Tyr Trp Asn Trp Ser Glu Asp Glu Gly Trp Gln Ser
675 680 685

cag ttt gat caa aaa gat ctg atc gaa aat gaa caa act ctg gat ctg 2112
Gln Phe Asp Gln Lys Asp Leu Ile Glu Asn Glu Gln Thr Leu Asp Leu
690 695 700

aaa gcg gac gaa acc ggc aag gtc agt ttt ccg gta gag tgg ggc gct 2160
Lys Ala Asp Glu Thr Gly Lys Val Ser Phe Pro Val Glu Trp Gly Ala
705 710 715 720

tat cgt ctg gaa gtc aaa gcg ccg aat gaa gcg gtc agt agt gtt cgt 2208
Tyr Arg Leu Glu Val Lys Ala Pro Asn Glu Ala Val Ser Ser Val Arg
725 730 735

ttc tgg gct ggc tat agc tgg cag gac aac agc gac ggt agc ggc gca 2256
Phe Trp Ala Gly Tyr Ser Trp Gln Asp Asn Ser Asp Gly Ser Gly Ala
740 745 750

gtg cga ccc gac cgt gtc acg ctg aaa ctg gat aaa gcc agt tat cgc 2304
Val Arg Pro Asp Arg Val Thr Leu Lys Leu Asp Lys Ala Ser Tyr Arg
755 760 765

cct ggc gac acc att aag ttg cat atc gcc gcg cca acg gcg ggt aaa 2352
Pro Gly Asp Thr Ile Lys Leu His Ile Ala Ala Pro Thr Ala Gly Lys
770 775 780

ggt tat gcg atg gtc gag tcc agt gaa ggg ccg ctg tgg tgg caa gag 2400
Gly Tyr Ala Met Val Glu Ser Ser Glu Gly Pro Leu Trp Trp Gln Glu

785 790 795 800

att gat gtt cgg gct caa ggg ctg gat ctg acg att ccg gtc gat aaa 2448
Ile Asp Val Arg Ala Gln Gly Leu Asp Leu Thr Ile Pro Val Asp Lys
805 810 815

acc tgg aat cgt cat gat ctg tat tta agt acg ctg gtg gta cgt cct 2496
Thr Trp Asn Arg His Asp Leu Tyr Leu Ser Thr Leu Val Val Arg Pro
820 825 830

ggc gat aaa tct cgc tcc gcg acg cca aaa cgc gcg gtt ggt gtg ttg 2544
Gly Asp Lys Ser Arg Ser Ala Thr Pro Lys Arg Ala Val Gly Val Leu
835 840 845

cat ctg ccg ctt ggc gat gaa aac cgt cgc ctc gat ctg gcg ctg gaa 2592
His Leu Pro Leu Gly Asp Glu Asn Arg Arg Leu Asp Leu Ala Leu Glu
850 855 860

aca cca gca aaa atg cgt ccc aat caa cca tta acc gtg aaa att aaa 2640
Thr Pro Ala Lys Met Arg Pro Asn Gln Pro Leu Thr Val Lys Ile Lys
865 870 875 880

gcc agc act aaa aat ggc gag aag cct aaa cag gtg aat gtg ctg gtg 2688
Ala Ser Thr Lys Asn Gly Glu Lys Pro Lys Gln Val Asn Val Leu Val
885 890 895

tct gcc gtt gat agt ggt gtg ctg aat att act gac tac gtc acg cca 2736
Ser Ala Val Asp Ser Gly Val Leu Asn Ile Thr Asp Tyr Val Thr Pro
900 905 910

gat ccg tgg cag gcg ttc ttt ggt cag aaa cgc tat ggc gca gac att 2784
Asp Pro Trp Gln Ala Phe Phe Gly Gln Lys Arg Tyr Gly Ala Asp Ile
915 920 925

tac gat att tac ggt cag gtt att gaa ggt cag ggg cgt ctg gca gct 2832
Tyr Asp Ile Tyr Gly Gln Val Ile Glu Gly Gln Gly Arg Leu Ala Ala
930 935 940

ctg cgt ttc ggt ggc gat ggt gat gag ctg aaa cgt ggt ggt aaa ccg 2880
Leu Arg Phe Gly Gly Asp Gly Asp Glu Leu Lys Arg Gly Gly Lys Pro
945 950 955 960

ccg gtc aat cac gtc aat att gtc gtg cag cag gcg ctg ccg gta acg 2928
Pro Val Asn His Val Asn Ile Val Val Gln Gln Ala Leu Pro Val Thr
965 970 975

ctc aac gaa cag ggc gaa ggc tcg gtt aca ctg ccg att ggc gat ttt 2976
Leu Asn Glu Gln Gly Glu Gly Ser Val Thr Leu Pro Ile Gly Asp Phe
980 985 990

aac ggt gaa ttg cgc gtc atg gcg caa gcc tgg acg gca gat gac ttc 3024
Asn Gly Glu Leu Arg Val Met Ala Gln Ala Trp Thr Ala Asp Asp Phe
995 1000 1005

ggt agc aac gaa agt aaa gtg ata gtt gcc gca ccg gtg att gct gaa 3072
Gly Ser Asn Glu Ser Lys Val Ile Val Ala Ala Pro Val Ile Ala Glu
1010 1015 1020

ctg aac atg ccg cgc ttt atg gcg agt ggc gat acc tcg cgt ctg acg 3120
Leu Asn Met Pro Arg Phe Met Ala Ser Gly Asp Thr Ser Arg Leu Thr
1025 1030 1035 1040

```
ctg gat atc act aat ctt acc gat aaa ccg caa aaa ctg aac gtt gcc    3168
Leu Asp Ile Thr Asn Leu Thr Asp Lys Pro Gln Lys Leu Asn Val Ala
                1045                1050                1055

ctg acc gcc agt ggt ttg ctt gaa ctg gtc agc gat tca ccc gca gcc    3216
Leu Thr Ala Ser Gly Leu Leu Glu Leu Val Ser Asp Ser Pro Ala Ala
            1060                1065                1070

gtt gaa tta gcg cca ggt gtg cgt act acg ctg ttt atc ccg gtg cga    3264
Val Glu Leu Ala Pro Gly Val Arg Thr Thr Leu Phe Ile Pro Val Arg
        1075                1080                1085

gca ttg ccg ggt tat ggc gat gga gaa att cag gcc acc att agc ggg    3312
Ala Leu Pro Gly Tyr Gly Asp Gly Glu Ile Gln Ala Thr Ile Ser Gly
    1090                1095                1100

tta gcg tta ccg ggt gaa acc gtt gcc gat cag cat aag cag tgg aaa    3360
Leu Ala Leu Pro Gly Glu Thr Val Ala Asp Gln His Lys Gln Trp Lys
1105                1110                1115                1120

atc ggc gtc cgt ccg gcg ttc ccg gca caa acg gtt aat tac ggt acg    3408
Ile Gly Val Arg Pro Ala Phe Pro Ala Gln Thr Val Asn Tyr Gly Thr
                1125                1130                1135

gcg tta cag cct ggt gag aca tgg gcg att ccg gcg gat gga ttg caa    3456
Ala Leu Gln Pro Gly Glu Thr Trp Ala Ile Pro Ala Asp Gly Leu Gln
            1140                1145                1150

aac ttc tcg cct gtt acg ctg gaa ggg caa ttg ttg ttg agc ggc aaa    3504
Asn Phe Ser Pro Val Thr Leu Glu Gly Gln Leu Leu Leu Ser Gly Lys
        1155                1160                1165

cca ccg ctg aac atc gca cgt tat atc aaa gag tta aaa gcg tat ccg    3552
Pro Pro Leu Asn Ile Ala Arg Tyr Ile Lys Glu Leu Lys Ala Tyr Pro
    1170                1175                1180

tac ggc tgt ctt gag caa acc gcc agc ggc ctg ttt ccg tca ctt tat    3600
Tyr Gly Cys Leu Glu Gln Thr Ala Ser Gly Leu Phe Pro Ser Leu Tyr
1185                1190                1195                1200

acc aac gca gcc caa ctg cag gcg ttg ggc atc aaa ggc gac agt gat    3648
Thr Asn Ala Ala Gln Leu Gln Ala Leu Gly Ile Lys Gly Asp Ser Asp
                1205                1210                1215

gag aaa cgc cgt gca tcg gtc gat atc ggc att tcc cgt ttg ctg caa    3696
Glu Lys Arg Arg Ala Ser Val Asp Ile Gly Ile Ser Arg Leu Leu Gln
            1220                1225                1230

atg caa cgt gat aac ggc ggc ttt gcg ctg tgg gat aaa aac ggt gac    3744
Met Gln Arg Asp Asn Gly Gly Phe Ala Leu Trp Asp Lys Asn Gly Asp
        1235                1240                1245

gaa gag tac tgg ctg acg gct tac gtg atg gat ttc ctg gtc cgc gca    3792
Glu Glu Tyr Trp Leu Thr Ala Tyr Val Met Asp Phe Leu Val Arg Ala
    1250                1255                1260

ggc gaa cag ggt tac agc gtg ccg aca gac gcc att aac cgg ggt aat    3840
Gly Glu Gln Gly Tyr Ser Val Pro Thr Asp Ala Ile Asn Arg Gly Asn
1265                1270                1275                1280
```

gag cgt ctg ctg cgc tat tta caa gat ccg ggc atg atg tcg atc ccg 3888
Glu Arg Leu Leu Arg Tyr Leu Gln Asp Pro Gly Met Met Ser Ile Pro
1285 1290 1295

tac gcg gat aat ctc aaa gcc agt aaa ttc gcc gta cag tct tac gct 3936
Tyr Ala Asp Asn Leu Lys Ala Ser Lys Phe Ala Val Gln Ser Tyr Ala
1300 1305 1310

gcg ctg gtg ttg gcc cgt cag caa aag gct ccg ctg ggt gcg ctg cgt 3984
Ala Leu Val Leu Ala Arg Gln Gln Lys Ala Pro Leu Gly Ala Leu Arg
1315 1320 1325

gaa atc tgg gag cat cgt gca gat gcc gct tct ggt tta ccg ctg ctg 4032
Glu Ile Trp Glu His Arg Ala Asp Ala Ala Ser Gly Leu Pro Leu Leu
1330 1335 1340

caa ctt ggc gtt gcg ctg aaa acc atg ggt gat gcg acg cgt ggt gaa 4080
Gln Leu Gly Val Ala Leu Lys Thr Met Gly Asp Ala Thr Arg Gly Glu
1345 1350 1355 1360

gaa gcg att gcg ctg gcg ctg aaa acg ccg cgt aat agt gat gag cgg 4128
Glu Ala Ile Ala Leu Ala Leu Lys Thr Pro Arg Asn Ser Asp Glu Arg
1365 1370 1375

ata tgg ctg ggt gat tac ggt agt tca ctg cgc gac aac gcg tta atg 4176
Ile Trp Leu Gly Asp Tyr Gly Ser Ser Leu Arg Asp Asn Ala Leu Met
1380 1385 1390

ctc tcc ttg ctg gaa gaa aat aaa ctg cta ccc gat gag cag tac act 4224
Leu Ser Leu Leu Glu Glu Asn Lys Leu Leu Pro Asp Glu Gln Tyr Thr
1395 1400 1405

ttg ctg aac aca ctt tcg cag cag gcg ttt ggt gaa cgc tgg cta tcg 4272
Leu Leu Asn Thr Leu Ser Gln Gln Ala Phe Gly Glu Arg Trp Leu Ser
1410 1415 1420

acg cag gaa agt aac gcg ttg ttc ctg gct gcc cgt acg att cag gat 4320
Thr Gln Glu Ser Asn Ala Leu Phe Leu Ala Ala Arg Thr Ile Gln Asp
1425 1430 1435 1440

tta ccc ggt aaa tgg cag gcg caa acc tct ttc tca gct gag cag ctg 4368
Leu Pro Gly Lys Trp Gln Ala Gln Thr Ser Phe Ser Ala Glu Gln Leu
1445 1450 1455

aca ggc gag aaa gcg caa aac agc aat ctg aat agc gat caa ctt gtc 4416
Thr Gly Glu Lys Ala Gln Asn Ser Asn Leu Asn Ser Asp Gln Leu Val
1460 1465 1470

acc ttg cag gtg agc aac agt ggc gat cag ccg tta tgg ttg cgt atg 4464
Thr Leu Gln Val Ser Asn Ser Gly Asp Gln Pro Leu Trp Leu Arg Met
1475 1480 1485

gat gcc agc ggt tat ccg caa tcc gca cct tta ccg gcg aac aat gtg 4512
Asp Ala Ser Gly Tyr Pro Gln Ser Ala Pro Leu Pro Ala Asn Asn Val
1490 1495 1500

ctg caa atc gag cgt cat att ctt ggt act gat ggt aag agc aaa tcg 4560
Leu Gln Ile Glu Arg His Ile Leu Gly Thr Asp Gly Lys Ser Lys Ser
1505 1510 1515 1520

ctg gac tcg tta cgt agc ggc gat ctg gtg ctg gtg tgg ttg cag gta 4608

Leu Asp Ser Leu Arg Ser Gly Asp Leu Val Leu Val Trp Leu Gln Val
1525 1530 1535

aaa gcc agt aac agc gtg ccg gat gcg tta gtc gtg gat ctg ctg cct 4656
Lys Ala Ser Asn Ser Val Pro Asp Ala Leu Val Val Asp Leu Leu Pro
1540 1545 1550

gcg ggt ctg gaa ctg gaa aac cag aat ctg gcg aac ggt agc gcc agc 4704
Ala Gly Leu Glu Leu Glu Asn Gln Asn Leu Ala Asn Gly Ser Ala Ser
1555 1560 1565

ctg gag caa agt ggt ggc gaa gtg cag aac tta ctg aac cag atg cag 4752
Leu Glu Gln Ser Gly Gly Glu Val Gln Asn Leu Leu Asn Gln Met Gln
1570 1575 1580

cag gcg agc att aag cac att gag ttc cgt gac gat cgc ttt gtg gcg 4800
Gln Ala Ser Ile Lys His Ile Glu Phe Arg Asp Asp Arg Phe Val Ala
1585 1590 1595 1600

gcg gtt gcc gtt gat gaa tac caa ccg gta acg ctg gtg tat ctg gcg 4848
Ala Val Ala Val Asp Glu Tyr Gln Pro Val Thr Leu Val Tyr Leu Ala
1605 1610 1615

cgg gcg gtg acg ccg gga acg tat cag gta ccg caa ccg atg gtg gaa 4896
Arg Ala Val Thr Pro Gly Thr Tyr Gln Val Pro Gln Pro Met Val Glu
1620 1625 1630

tca atg tat gtt ccc caa tgg cgg gcg acc ggc gcg gct gaa gat ctg 4944
Ser Met Tyr Val Pro Gln Trp Arg Ala Thr Gly Ala Ala Glu Asp Leu
1635 1640 1645

ctg att gtc aga ccg taa 4962
Leu Ile Val Arg Pro *
1650

<210> 210
<211> 795
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(795)

<400> 210
atg gat ctg ttg tac cgg gta aaa acg ctt tgg gcc gcg ctg cgc ggt 48
Met Asp Leu Leu Tyr Arg Val Lys Thr Leu Trp Ala Ala Leu Arg Gly
1 5 10 15

aat cat tac acc tgg cct gcc atc gat atc acc ctc ccc ggc aat cgc 96
Asn His Tyr Thr Trp Pro Ala Ile Asp Ile Thr Leu Pro Gly Asn Arg
20 25 30

cat ttt cat ctg att ggc agt att cat atg ggt agc cac gat atg gct 144
His Phe His Leu Ile Gly Ser Ile His Met Gly Ser His Asp Met Ala
35 40 45

ccc ctg ccc acc cgt ttg ctc aaa aag ctc aaa aac gcc gat gcg ctg 192
Pro Leu Pro Thr Arg Leu Leu Lys Lys Leu Lys Asn Ala Asp Ala Leu
50 55 60

```
atc gtc gag gcg gat gtt tcc acc agc gat acg cct ttt gct aat ttg    240
Ile Val Glu Ala Asp Val Ser Thr Ser Asp Thr Pro Phe Ala Asn Leu
 65                  70                  75                  80

cct gcc tgc gag gcg ctg gaa gag cgc att agc gag gag caa ctc caa    288
Pro Ala Cys Glu Ala Leu Glu Glu Arg Ile Ser Glu Glu Gln Leu Gln
                 85                  90                  95

aac ctg cag cac att agc cag gag atg ggc att tct ccc tca ctc ttt    336
Asn Leu Gln His Ile Ser Gln Glu Met Gly Ile Ser Pro Ser Leu Phe
            100                 105                 110

tct acc caa ccg ctg tgg caa atc gcg atg gtt ctt cag gcg acg cag    384
Ser Thr Gln Pro Leu Trp Gln Ile Ala Met Val Leu Gln Ala Thr Gln
        115                 120                 125

gca caa aaa ctg ggg ctg cgg gca gaa tac ggt atc gat tac cag cta    432
Ala Gln Lys Leu Gly Leu Arg Ala Glu Tyr Gly Ile Asp Tyr Gln Leu
    130                 135                 140

ttg cag gcg gcg aag caa caa cat aaa ccc gtg att gaa ctg gaa ggg    480
Leu Gln Ala Ala Lys Gln Gln His Lys Pro Val Ile Glu Leu Glu Gly
145                 150                 155                 160

gct gaa aac cag att gcc atg ttg ctc cag ctc cct gac aaa gga ctg    528
Ala Glu Asn Gln Ile Ala Met Leu Leu Gln Leu Pro Asp Lys Gly Leu
                165                 170                 175

gcg ctg ctg gac gat acg ctg acc cac tgg cat acc aac gca cgg ttg    576
Ala Leu Leu Asp Asp Thr Leu Thr His Trp His Thr Asn Ala Arg Leu
            180                 185                 190

ctg caa caa atg atg agc tgg tgg ctg aat gca ccg ccg caa aat aat    624
Leu Gln Gln Met Met Ser Trp Trp Leu Asn Ala Pro Pro Gln Asn Asn
        195                 200                 205

gat ata acg ctg ccc aat acg ttc agt cag tcg ctg tac gat gtg ctg    672
Asp Ile Thr Leu Pro Asn Thr Phe Ser Gln Ser Leu Tyr Asp Val Leu
    210                 215                 220

atg cat cag cga aat ctc gcc tgg cgg gat aaa tta cgc gcc atg ccg    720
Met His Gln Arg Asn Leu Ala Trp Arg Asp Lys Leu Arg Ala Met Pro
225                 230                 235                 240

ccg ggg cga tat gtg gtc gcg gtc ggt gca cta cac ctg tat gga gaa    768
Pro Gly Arg Tyr Val Val Ala Val Gly Ala Leu His Leu Tyr Gly Glu
                245                 250                 255

ggg aat ttg ccg caa atg ttg cgc taa                                795
Gly Asn Leu Pro Gln Met Leu Arg  *
            260
```

<210> 211
<211> 969
<212> DNA
<213> Escherichia coli

<220>
<221> CDS

<222> (1)...(969)

<400> 211
atg gat aat ttt ctt gct ctg acc tta acg ggt aaa aaa ccg gtt atc 48
Met Asp Asn Phe Leu Ala Leu Thr Leu Thr Gly Lys Lys Pro Val Ile
1 5 10 15

acc gag cga gaa atc aac ggc gtt cgc tgg cgc tgg ctg ggc gat ggt 96
Thr Glu Arg Glu Ile Asn Gly Val Arg Trp Arg Trp Leu Gly Asp Gly
20 25 30

gtg ctg gaa ctg acg cca tta acg cca ccg caa ggc gca ctg gtg att 144
Val Leu Glu Leu Thr Pro Leu Thr Pro Pro Gln Gly Ala Leu Val Ile
35 40 45

tca gcg gga ata cac ggt aat gag acg gca cct gtg gag atg ctg gac 192
Ser Ala Gly Ile His Gly Asn Glu Thr Ala Pro Val Glu Met Leu Asp
50 55 60

gcg ttg ctt ggc gcg ata tct cac ggc gag atc ccg tta cgt tgg cgg 240
Ala Leu Leu Gly Ala Ile Ser His Gly Glu Ile Pro Leu Arg Trp Arg
65 70 75 80

ttg ctg gtg atc ctc ggg aat cct cct gcg ctg aag caa ggg aaa cgt 288
Leu Leu Val Ile Leu Gly Asn Pro Pro Ala Leu Lys Gln Gly Lys Arg
85 90 95

tat tgc cat agc gat atg aat cga atg ttt ggc ggt cgt tgg cag cta 336
Tyr Cys His Ser Asp Met Asn Arg Met Phe Gly Gly Arg Trp Gln Leu
100 105 110

ttt gct gaa agc gga gaa acc tgt cgg gcg cgc gaa ctg gaa cag tgc 384
Phe Ala Glu Ser Gly Glu Thr Cys Arg Ala Arg Glu Leu Glu Gln Cys
115 120 125

ctg gaa gat ttt tat gac cag ggc aaa gaa tct gtg cgc tgg cac ctt 432
Leu Glu Asp Phe Tyr Asp Gln Gly Lys Glu Ser Val Arg Trp His Leu
130 135 140

gat cta cat acc gca att cgt ggc tcc ttg cat ccg cag ttc ggt gta 480
Asp Leu His Thr Ala Ile Arg Gly Ser Leu His Pro Gln Phe Gly Val
145 150 155 160

tta ccg caa cgc gac att ccc tgg gac gag aaa ttt ctg acg tgg ctg 528
Leu Pro Gln Arg Asp Ile Pro Trp Asp Glu Lys Phe Leu Thr Trp Leu
165 170 175

ggt gcg gcg ggg ctg gag gcg ctg gtg ttc cat cag gaa cct ggt ggt 576
Gly Ala Ala Gly Leu Glu Ala Leu Val Phe His Gln Glu Pro Gly Gly
180 185 190

acg ttt acc cat ttc agc gcc aga cat ttt ggc gcg ctg gcc tgt acg 624
Thr Phe Thr His Phe Ser Ala Arg His Phe Gly Ala Leu Ala Cys Thr
195 200 205

ctg gaa ctt ggc aaa gcg ttg ccc ttt ggg caa aac gat ctt cgc cag 672
Leu Glu Leu Gly Lys Ala Leu Pro Phe Gly Gln Asn Asp Leu Arg Gln
210 215 220

ttt gca gta act gcc agc gca att gct gcg ctg cta tct ggt gag agt 720
Phe Ala Val Thr Ala Ser Ala Ile Ala Ala Leu Leu Ser Gly Glu Ser

225 230 235 240

gtc ggt atc gtg aga aca ccg ccg ctc cgt tat cgg gtg gtt tcg caa 768
Val Gly Ile Val Arg Thr Pro Pro Leu Arg Tyr Arg Val Val Ser Gln
245 250 255

att act cgc cac tcg ccg tcc ttc gaa atg cat atg gca agt gac acg 816
Ile Thr Arg His Ser Pro Ser Phe Glu Met His Met Ala Ser Asp Thr
260 265 270

ctg aat ttt atg ccg ttt gag aaa gga aca ttg ctg gcg cag gac gga 864
Leu Asn Phe Met Pro Phe Glu Lys Gly Thr Leu Leu Ala Gln Asp Gly
275 280 285

gag gaa cgt ttt acc gta acc cat gat gta gag tat gtg tta ttc cct 912
Glu Glu Arg Phe Thr Val Thr His Asp Val Glu Tyr Val Leu Phe Pro
290 295 300

aat ccg ttg gta gcg ttg gga tta cgc gcg gga tta atg ctc gaa aaa 960
Asn Pro Leu Val Ala Leu Gly Leu Arg Ala Gly Leu Met Leu Glu Lys
305 310 315 320

ata agc taa 969
Ile Ser *

<210> 212
<211> 1344
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1344)

<400> 212
atg aac gcc tgg gaa gtc aat ttc gac ggg ctg gta ggg ctg acg cat 48
Met Asn Ala Trp Glu Val Asn Phe Asp Gly Leu Val Gly Leu Thr His
1 5 10 15

cat tac gcg ggc ctg tcg ttt ggt aat gaa gcc tct acc cgt cac cgt 96
His Tyr Ala Gly Leu Ser Phe Gly Asn Glu Ala Ser Thr Arg His Arg
20 25 30

ttt cag gtg tct aac ccg cga ctg gcg gcg aag cag ggc tta ctg aaa 144
Phe Gln Val Ser Asn Pro Arg Leu Ala Ala Lys Gln Gly Leu Leu Lys
35 40 45

atg aaa gcc ctt gcc gat gcg gga ttc ccc cag gcc gtg atc ccg ccg 192
Met Lys Ala Leu Ala Asp Ala Gly Phe Pro Gln Ala Val Ile Pro Pro
50 55 60

cac gag cgt ccg ttt att ccg gtg ctg cgt cag ttg gga ttc agt ggt 240
His Glu Arg Pro Phe Ile Pro Val Leu Arg Gln Leu Gly Phe Ser Gly
65 70 75 80

agc gat gag cag gta ctg gaa aaa gtt gca cgc cag gca ccg cac tgg 288
Ser Asp Glu Gln Val Leu Glu Lys Val Ala Arg Gln Ala Pro His Trp
85 90 95

```
ctt tcc agc gtc agt tcc gct tcg cca atg tgg gta gcc aat gcg gca   336
Leu Ser Ser Val Ser Ser Ala Ser Pro Met Trp Val Ala Asn Ala Ala
            100                 105                 110

acg atc gcg cca tct gcc gat acg ctg gat ggc aaa gtg cat ctc acc   384
Thr Ile Ala Pro Ser Ala Asp Thr Leu Asp Gly Lys Val His Leu Thr
        115                 120                 125

gtt gcc aac ctg aac aat aaa ttt cac cgt tcg ctg gaa gcg ccc gtc   432
Val Ala Asn Leu Asn Asn Lys Phe His Arg Ser Leu Glu Ala Pro Val
    130                 135                 140

act gaa tcg ctg tta aaa gcg att ttt aac gac gaa gag aaa ttt agc   480
Thr Glu Ser Leu Leu Lys Ala Ile Phe Asn Asp Glu Glu Lys Phe Ser
145                 150                 155                 160

gtc cat tcg gcg ttg cca cag gta gcg ttg ctc ggt gat gag ggg gcg   528
Val His Ser Ala Leu Pro Gln Val Ala Leu Leu Gly Asp Glu Gly Ala
                165                 170                 175

gca aac cac aat cgt ctc ggc ggt cat tac ggt gaa ccg ggt atg caa   576
Ala Asn His Asn Arg Leu Gly Gly His Tyr Gly Glu Pro Gly Met Gln
            180                 185                 190

ctt ttt gtc tac ggg cga gaa gaa ggc aat gat acc cgg cct tcc cgt   624
Leu Phe Val Tyr Gly Arg Glu Glu Gly Asn Asp Thr Arg Pro Ser Arg
        195                 200                 205

tat ccg gcg cga cag act cgc gaa gcc agc gag gcg gtg gca agg ctg   672
Tyr Pro Ala Arg Gln Thr Arg Glu Ala Ser Glu Ala Val Ala Arg Leu
    210                 215                 220

aat cag gtg aat ccc caa cag gtg att ttc gcc cag caa aac ccg gac   720
Asn Gln Val Asn Pro Gln Gln Val Ile Phe Ala Gln Gln Asn Pro Asp
225                 230                 235                 240

gtt atc gac cag ggc gtt ttt cat aat gac gtg att gcc gtg agt aac   768
Val Ile Asp Gln Gly Val Phe His Asn Asp Val Ile Ala Val Ser Asn
                245                 250                 255

cgc cag gtg ctg ttt tgc cac caa cag gcg ttc gct cgc cag tca cag   816
Arg Gln Val Leu Phe Cys His Gln Gln Ala Phe Ala Arg Gln Ser Gln
            260                 265                 270

tta ctg gca aac ctg cgt gcg cgg gtc aat ggt ttt atg gcg ata gaa   864
Leu Leu Ala Asn Leu Arg Ala Arg Val Asn Gly Phe Met Ala Ile Glu
        275                 280                 285

gtt ccg gca act cag gtt tcc gtg tct gat acg gtg tct acc tat ctg   912
Val Pro Ala Thr Gln Val Ser Val Ser Asp Thr Val Ser Thr Tyr Leu
    290                 295                 300

ttt aac agc caa ctg ctg agc cgc gat gat ggt tcc atg atg ttg gtg   960
Phe Asn Ser Gln Leu Leu Ser Arg Asp Asp Gly Ser Met Met Leu Val
305                 310                 315                 320

ctg cct cag gag tgt cgg gaa cac gcc gga gta tgg ggt tat ctc aat   1008
Leu Pro Gln Glu Cys Arg Glu His Ala Gly Val Trp Gly Tyr Leu Asn
                325                 330                 335

gaa ctc ctt gcc gct gac aac ccg att agc gaa cta aaa gtc ttt gat   1056
```

```
Glu Leu Leu Ala Ala Asp Asn Pro Ile Ser Glu Leu Lys Val Phe Asp
            340                 345                 350

tta cgt gaa agc atg gcg aat ggc ggc ccg gcg tgc ctg cgg ttg     1104
Leu Arg Glu Ser Met Ala Asn Gly Gly Gly Pro Ala Cys Leu Arg Leu
        355                 360                 365

cgg gtg gta ttg aca gaa gaa gaa cgc cgg gcg gtg aat ccg gcg gtg     1152
Arg Val Val Leu Thr Glu Glu Glu Arg Arg Ala Val Asn Pro Ala Val
    370                 375                 380

atg atg aac gat acg ctg ttt aat gcg ctc aat gac tgg gtg gat cgt     1200
Met Met Asn Asp Thr Leu Phe Asn Ala Leu Asn Asp Trp Val Asp Arg
385                 390                 395                 400

tac tac cgc gat cgc ctt act gct gcc gat ctg gcc gac ccg caa ttg     1248
Tyr Tyr Arg Asp Arg Leu Thr Ala Ala Asp Leu Ala Asp Pro Gln Leu
                405                 410                 415

ctg cgc gaa ggg cgg gaa gca ctg gat gta ttg agc caa tta ctg aat     1296
Leu Arg Glu Gly Arg Glu Ala Leu Asp Val Leu Ser Gln Leu Leu Asn
            420                 425                 430

ctc ggt tcg gtt tat ccg ttc cag cgc gag gga ggg ggc aat gga taa     1344
Leu Gly Ser Val Tyr Pro Phe Gln Arg Glu Gly Gly Gly Asn Gly  *
        435                 440                 445
```

<210> 213
<211> 1479
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1479)

<400> 213

```
atg act tta tgg att aac ggt gac tgg ata acg ggc cag ggc gca tcg     48
Met Thr Leu Trp Ile Asn Gly Asp Trp Ile Thr Gly Gln Gly Ala Ser
 1               5                  10                  15

cgt gtg aag cgt aat ccg gta tcg ggc gag gtg tta tgg caa ggc aat     96
Arg Val Lys Arg Asn Pro Val Ser Gly Glu Val Leu Trp Gln Gly Asn
            20                  25                  30

gat gcc gat gcc gct cag gtc gag cag gct tgt cgg gca gcc cgt gcg     144
Asp Ala Asp Ala Ala Gln Val Glu Gln Ala Cys Arg Ala Ala Arg Ala
        35                  40                  45

gcg ttt ccg cgc tgg gcg cgg ctt tca ttt gct gaa cgt cat gcc gtt     192
Ala Phe Pro Arg Trp Ala Arg Leu Ser Phe Ala Glu Arg His Ala Val
    50                  55                  60

gtc gaa cgc ttt gcc gca ctg ctg gaa agc aat aaa gcc gaa tta acc     240
Val Glu Arg Phe Ala Ala Leu Leu Glu Ser Asn Lys Ala Glu Leu Thr
65                  70                  75                  80

gcg att att gcc aga gaa acg ggt aag ccg cgc tgg gaa gcg gca acc     288
Ala Ile Ile Ala Arg Glu Thr Gly Lys Pro Arg Trp Glu Ala Ala Thr
```

```
                85                  90                  95

gaa gtg acg gcg atg atc aat aaa atc gcg ata tca att aag gcg tat    336
Glu Val Thr Ala Met Ile Asn Lys Ile Ala Ile Ser Ile Lys Ala Tyr
            100                 105                 110

cac gtt cgt acc ggc gag cag cgt agt gaa atg ccg gac ggc gcg gcg    384
His Val Arg Thr Gly Glu Gln Arg Ser Glu Met Pro Asp Gly Ala Ala
        115                 120                 125

agc ctg cga cat cgc ccg cac ggc gtg ctg gcg gtg ttt ggg ccg tat    432
Ser Leu Arg His Arg Pro His Gly Val Leu Ala Val Phe Gly Pro Tyr
    130                 135                 140

aat ttc cct ggt cat ttg ccg aac gga cat atc gtt ccg gca ttg ctg    480
Asn Phe Pro Gly His Leu Pro Asn Gly His Ile Val Pro Ala Leu Leu
145                 150                 155                 160

gca ggt aac acc att atc ttt aaa ccc agc gaa ctg aca ccg tgg agt    528
Ala Gly Asn Thr Ile Ile Phe Lys Pro Ser Glu Leu Thr Pro Trp Ser
            165                 170                 175

ggc gaa gcg gta atg cgt tta tgg cag cag gct ggc ttg ccg cca ggc    576
Gly Glu Ala Val Met Arg Leu Trp Gln Gln Ala Gly Leu Pro Pro Gly
            180                 185                 190

gtg ctg aac ctg gtg cag ggc ggg cgt gaa acg ggt cag gcg ctg agt    624
Val Leu Asn Leu Val Gln Gly Gly Arg Glu Thr Gly Gln Ala Leu Ser
        195                 200                 205

gcg ctg gag gat ctc gac ggt ttg ctg ttt acc ggt agc gcc aat aca    672
Ala Leu Glu Asp Leu Asp Gly Leu Leu Phe Thr Gly Ser Ala Asn Thr
    210                 215                 220

ggc tac cag ttg cat cgc cag ctc tcc ggt cag ccg gag aaa att ctc    720
Gly Tyr Gln Leu His Arg Gln Leu Ser Gly Gln Pro Glu Lys Ile Leu
225                 230                 235                 240

gcc ctt gag atg ggc ggt aat aat ccg cta att atc gat gag gtg gcg    768
Ala Leu Glu Met Gly Gly Asn Asn Pro Leu Ile Ile Asp Glu Val Ala
            245                 250                 255

gat atc gac gcg gct gtc cat ctg acc att cag tcg gcg ttt gtc aca    816
Asp Ile Asp Ala Ala Val His Leu Thr Ile Gln Ser Ala Phe Val Thr
            260                 265                 270

gcc ggt caa cgc tgc acc tgc gcc cgc cgt tta ttg ctg aaa agc ggg    864
Ala Gly Gln Arg Cys Thr Cys Ala Arg Arg Leu Leu Leu Lys Ser Gly
        275                 280                 285

gcg cag ggc gat gcg ttt ctt gct cgt ctg gtt gcc gtc agc cag cga    912
Ala Gln Gly Asp Ala Phe Leu Ala Arg Leu Val Ala Val Ser Gln Arg
    290                 295                 300

tta acg ccg ggc aac tgg gat gac gaa ccg cag ccg ttt att ggc ggg    960
Leu Thr Pro Gly Asn Trp Asp Asp Glu Pro Gln Pro Phe Ile Gly Gly
305                 310                 315                 320

ctg att tct gaa cag gcc gca cag cag gtg gtt act gca tgg cag caa   1008
Leu Ile Ser Glu Gln Ala Ala Gln Gln Val Val Thr Ala Trp Gln Gln
            325                 330                 335
```

ctg gaa gcg atg ggc gga cga ccc ctg ctt gcg ccg cgc tta tta caa 1056
Leu Glu Ala Met Gly Gly Arg Pro Leu Leu Ala Pro Arg Leu Leu Gln
340 345 350

gca ggg aca tcg ttg ctg acg ccg ggg atc att gaa atg aca ggc gtt 1104
Ala Gly Thr Ser Leu Leu Thr Pro Gly Ile Ile Glu Met Thr Gly Val
355 360 365

gct ggc gta cca gat gaa gag gtg ttc gga ccg tta ttg cgc gtc tgg 1152
Ala Gly Val Pro Asp Glu Glu Val Phe Gly Pro Leu Leu Arg Val Trp
370 375 380

cgt tat gat act ttc gat gaa gcg att cga atg gcg aat aac act cgc 1200
Arg Tyr Asp Thr Phe Asp Glu Ala Ile Arg Met Ala Asn Asn Thr Arg
385 390 395 400

ttc gga ctc tct tgc ggt ctg gtt tcc ccc gag cgg gaa aag ttc gat 1248
Phe Gly Leu Ser Cys Gly Leu Val Ser Pro Glu Arg Glu Lys Phe Asp
405 410 415

caa ctg ttg ctg gag gcg cgg gcg ggg att gtt aac tgg aac aaa ccg 1296
Gln Leu Leu Leu Glu Ala Arg Ala Gly Ile Val Asn Trp Asn Lys Pro
420 425 430

ctt acc ggt gct gcc agt acc gcg cca ttc ggc ggc att ggt gct tcc 1344
Leu Thr Gly Ala Ala Ser Thr Ala Pro Phe Gly Gly Ile Gly Ala Ser
435 440 445

ggt aac cat cgc ccc agc gcc tgg tat gcc gca gat tac tgc gca tgg 1392
Gly Asn His Arg Pro Ser Ala Trp Tyr Ala Ala Asp Tyr Cys Ala Trp
450 455 460

ccg atg gcg agc ctg gag tcg gac tcg tta aca ttg ccc gcc acg ctt 1440
Pro Met Ala Ser Leu Glu Ser Asp Ser Leu Thr Leu Pro Ala Thr Leu
465 470 475 480

aac ccc ggg ctg gat ttt tcc gat gag gtg gtg cga tga 1479
Asn Pro Gly Leu Asp Phe Ser Asp Glu Val Val Arg *
485 490

<210> 214
<211> 1035
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1035)

<400> 214
atg atg gtc atc cgt ccc gtt gag cga tca gat gtc tcg gcg ctg atg 48
Met Met Val Ile Arg Pro Val Glu Arg Ser Asp Val Ser Ala Leu Met
1 5 10 15

cag ctt gcc agc aaa acg ggc ggc ggc ctg acg tcg ctt ccc gcc aat 96
Gln Leu Ala Ser Lys Thr Gly Gly Gly Leu Thr Ser Leu Pro Ala Asn
20 25 30

gaa gcc acg ctt tcg gcg cgt atc gaa agg gca atc aaa acc tgg caa 144

```
Glu Ala Thr Leu Ser Ala Arg Ile Glu Arg Ala Ile Lys Thr Trp Gln
        35                  40                  45

ggc gaa ctg ccc aaa agt gag cag ggc tat gtg ttc gtg ctg gaa gat      192
Gly Glu Leu Pro Lys Ser Glu Gln Gly Tyr Val Phe Val Leu Glu Asp
    50                  55                  60

agc gag aca ggc acc gtg gcg ggg att tgt gcc att gag gtg gcg gtt      240
Ser Glu Thr Gly Thr Val Ala Gly Ile Cys Ala Ile Glu Val Ala Val
65                  70                  75                  80

ggg ctg aac gat ccc tgg tac aac tat cgc gtc ggc acg ttg gtt cac      288
Gly Leu Asn Asp Pro Trp Tyr Asn Tyr Arg Val Gly Thr Leu Val His
                85                  90                  95

gcc tca aaa gag ctg aat gtc tat aac gca ttg ccg acg ctg ttt ctc      336
Ala Ser Lys Glu Leu Asn Val Tyr Asn Ala Leu Pro Thr Leu Phe Leu
            100                 105                 110

agt aac gat cac acc ggc agc agc gag ctg tgc acg ctg ttt ctc gac      384
Ser Asn Asp His Thr Gly Ser Ser Glu Leu Cys Thr Leu Phe Leu Asp
        115                 120                 125

ccg gac tgg cgc aaa gag ggc aac ggc tat ttg ctg tcg aaa tcg cgc      432
Pro Asp Trp Arg Lys Glu Gly Asn Gly Tyr Leu Leu Ser Lys Ser Arg
    130                 135                 140

ttt atg ttt atg gcg gct ttt cgc gac aag ttt aat gac aaa gtg gtt      480
Phe Met Phe Met Ala Ala Phe Arg Asp Lys Phe Asn Asp Lys Val Val
145                 150                 155                 160

gct gaa atg cgc ggg gtg att gac gaa cac ggc tat tca ccg ttc tgg      528
Ala Glu Met Arg Gly Val Ile Asp Glu His Gly Tyr Ser Pro Phe Trp
                165                 170                 175

caa agc ctc ggt aaa cgc ttc ttt tcg atg gat ttt agc cgc gcc gat      576
Gln Ser Leu Gly Lys Arg Phe Phe Ser Met Asp Phe Ser Arg Ala Asp
            180                 185                 190

ttt ctc tgc ggc acc ggg caa aag gca ttt att gca gaa ctg atg ccg      624
Phe Leu Cys Gly Thr Gly Gln Lys Ala Phe Ile Ala Glu Leu Met Pro
        195                 200                 205

aaa cat ccg atc tat acc cac ttt tta tcc cag gaa gcc cag gac gtc      672
Lys His Pro Ile Tyr Thr His Phe Leu Ser Gln Glu Ala Gln Asp Val
    210                 215                 220

atc ggt cag gta cat ccg caa acc gcg cct gcc cgc gcg gtg ctg gag      720
Ile Gly Gln Val His Pro Gln Thr Ala Pro Ala Arg Ala Val Leu Glu
225                 230                 235                 240

aaa gaa ggt ttt cgc tac cgt aac tat atc gac atc ttt gac ggt ggg      768
Lys Glu Gly Phe Arg Tyr Arg Asn Tyr Ile Asp Ile Phe Asp Gly Gly
                245                 250                 255

ccg acg ctt gag tgt gac atc gac cgc gtg cgc gcc atc cgt aaa agt      816
Pro Thr Leu Glu Cys Asp Ile Asp Arg Val Arg Ala Ile Arg Lys Ser
            260                 265                 270

cgg ctg gtg gaa gtg gca gaa ggg cag cct gcg cag ggc gat ttc cca      864
Arg Leu Val Glu Val Ala Glu Gly Gln Pro Ala Gln Gly Asp Phe Pro
```

```
        275                 280                 285

gcc tgc ctg gtc gcc aat gaa aat tat cac cat ttc cgc gtg gtg ctg    912
Ala Cys Leu Val Ala Asn Glu Asn Tyr His His Phe Arg Val Val Leu
    290                 295                 300

gtg cgt acc gat ccg gca acc gag cgt ttg att tta acc gcc gca caa    960
Val Arg Thr Asp Pro Ala Thr Glu Arg Leu Ile Leu Thr Ala Ala Gln
305                 310                 315                 320

ctg gat gcc ctc aaa tgc cac gcc ggg gat cgc gtt cgt ctg gtg cgc   1008
Leu Asp Ala Leu Lys Cys His Ala Gly Asp Arg Val Arg Leu Val Arg
                325                 330                 335

ctg tgc gca gag gag aaa aca gca tga                               1035
Leu Cys Ala Glu Glu Lys Thr Ala  *
            340


<210> 215
<211> 1221
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1221)

<400> 215
atg tct cag cca att acg cgt gaa aac ttt gat gaa tgg atg ata cct     48
Met Ser Gln Pro Ile Thr Arg Glu Asn Phe Asp Glu Trp Met Ile Pro
 1               5                  10                  15

gtt tac gct ccg gca ccc ttt ata ccg gta cgt ggc gaa ggt tcg cgc     96
Val Tyr Ala Pro Ala Pro Phe Ile Pro Val Arg Gly Glu Gly Ser Arg
            20                  25                  30

ttg tgg gat cag cag ggg aaa gag tat atc gac ttc gcg ggt ggc att    144
Leu Trp Asp Gln Gln Gly Lys Glu Tyr Ile Asp Phe Ala Gly Gly Ile
        35                  40                  45

gcg gtg aac gcg ctg ggc cat gcg cat ccg gaa ctg cgt gaa gcg ctg    192
Ala Val Asn Ala Leu Gly His Ala His Pro Glu Leu Arg Glu Ala Leu
    50                  55                  60

aac gaa cag gcg agt aag ttc tgg cat acc ggc aac ggt tac acc aac    240
Asn Glu Gln Ala Ser Lys Phe Trp His Thr Gly Asn Gly Tyr Thr Asn
65                  70                  75                  80

gag ccg gta ctg cga ctg gcg aaa aaa ttg atc gac gcc acg ttt gcc    288
Glu Pro Val Leu Arg Leu Ala Lys Lys Leu Ile Asp Ala Thr Phe Ala
                85                  90                  95

gat cgc gtc ttc ttt tgt aac tcc ggt gcg gaa gcc aac gaa gcg gcg    336
Asp Arg Val Phe Phe Cys Asn Ser Gly Ala Glu Ala Asn Glu Ala Ala
            100                 105                 110

cta aaa ctg gcg cgt aaa ttc gct cac gac cgc tac ggc agc cat aag    384
Leu Lys Leu Ala Arg Lys Phe Ala His Asp Arg Tyr Gly Ser His Lys
        115                 120                 125
```

agc ggc atc gtg gcg ttc aaa aat gcg ttt cat ggt cgc acg ctg ttt 432
Ser Gly Ile Val Ala Phe Lys Asn Ala Phe His Gly Arg Thr Leu Phe
130 135 140

act gtc agt gcg ggt ggg cag cca gcc tat tca cag gat ttt gcg cca 480
Thr Val Ser Ala Gly Gly Gln Pro Ala Tyr Ser Gln Asp Phe Ala Pro
145 150 155 160

ctg ccg gcg gat att cgt cat gct gca tat aac gat att aac tct gcc 528
Leu Pro Ala Asp Ile Arg His Ala Ala Tyr Asn Asp Ile Asn Ser Ala
165 170 175

agc gcg ctg att gac gac tct acc tgt gcg gtg att gtc gaa ccc atc 576
Ser Ala Leu Ile Asp Asp Ser Thr Cys Ala Val Ile Val Glu Pro Ile
180 185 190

cag ggg gaa ggc ggt gtg gtg cca gcc agc aac gcg ttt tta caa ggt 624
Gln Gly Glu Gly Gly Val Val Pro Ala Ser Asn Ala Phe Leu Gln Gly
195 200 205

ctg cgt gaa ttg tgt aac cgc cac aat gcg ctg ttg att ttt gat gaa 672
Leu Arg Glu Leu Cys Asn Arg His Asn Ala Leu Leu Ile Phe Asp Glu
210 215 220

gta caa acc ggc gtc ggg cgc acc ggg gaa ctg tat gcc tat atg cac 720
Val Gln Thr Gly Val Gly Arg Thr Gly Glu Leu Tyr Ala Tyr Met His
225 230 235 240

tac ggc gtg acg cct gat ctg tta act acc gcc aaa gcg ctg ggc ggc 768
Tyr Gly Val Thr Pro Asp Leu Leu Thr Thr Ala Lys Ala Leu Gly Gly
245 250 255

ggt ttc ccg gtc ggt gcg ttg ttg gca acc gaa gag tgc gcc cgc gtg 816
Gly Phe Pro Val Gly Ala Leu Leu Ala Thr Glu Glu Cys Ala Arg Val
260 265 270

atg acc gtt ggc act cat ggc acc acc tat ggc ggt aat ccg ctg gcc 864
Met Thr Val Gly Thr His Gly Thr Thr Tyr Gly Gly Asn Pro Leu Ala
275 280 285

tcg gcg gtg gca ggc aaa gtg ctg gag ctc atc aac aca cca gag atg 912
Ser Ala Val Ala Gly Lys Val Leu Glu Leu Ile Asn Thr Pro Glu Met
290 295 300

ctt aat ggc gtt aaa cag cgt cac gac tgg ttt gtt gag cgt ctt aat 960
Leu Asn Gly Val Lys Gln Arg His Asp Trp Phe Val Glu Arg Leu Asn
305 310 315 320

act att aat cac cgc tat ggt ttg ttc agt gaa gtt cgc ggc tta ggt 1008
Thr Ile Asn His Arg Tyr Gly Leu Phe Ser Glu Val Arg Gly Leu Gly
325 330 335

ttg ctg att ggc tgt gta ctg aat gcc gat tac gcc ggg caa gcg aaa 1056
Leu Leu Ile Gly Cys Val Leu Asn Ala Asp Tyr Ala Gly Gln Ala Lys
340 345 350

cag atc tct cag gaa gcg gcg aaa gca ggc gtg atg gta ctg att gcg 1104
Gln Ile Ser Gln Glu Ala Ala Lys Ala Gly Val Met Val Leu Ile Ala
355 360 365

ggt ggc aac gtg gtg cgt ttt gcg cct gcg ctc aat gtc agc gaa gaa 1152

```
Gly Gly Asn Val Val Arg Phe Ala Pro Ala Leu Asn Val Ser Glu Glu
    370                 375                 380

gag gtg acg acc gga ctg gat cgc ttt gca gct gct tgc gaa cac ttt    1200
Glu Val Thr Thr Gly Leu Asp Arg Phe Ala Ala Ala Cys Glu His Phe
385                 390                 395                 400

gtt agc cga ggt tca tca tga                                        1221
Val Ser Arg Gly Ser Ser  *
                405
```

<210> 216
<211> 3147
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(3147)

<400> 216

```
atg aaa att ctc agc ctg cgc ctg aaa aac ctg aac tca tta aaa ggc    48
Met Lys Ile Leu Ser Leu Arg Leu Lys Asn Leu Asn Ser Leu Lys Gly
 1               5                  10                  15

gaa tgg aag att gat ttc acc cgc gag ccg ttc gcc agc aac ggg ctg    96
Glu Trp Lys Ile Asp Phe Thr Arg Glu Pro Phe Ala Ser Asn Gly Leu
            20                  25                  30

ttt gct att acc ggc cca aca ggt gcg ggg aaa acc acc ctg ctg gac    144
Phe Ala Ile Thr Gly Pro Thr Gly Ala Gly Lys Thr Thr Leu Leu Asp
        35                  40                  45

gcc att tgt ctg gcg ctg tat cac gaa act ccg cgt ctc tct aac gtt    192
Ala Ile Cys Leu Ala Leu Tyr His Glu Thr Pro Arg Leu Ser Asn Val
    50                  55                  60

tca caa tcg caa aat gat ctc atg acc cgc gat acc gcc gaa tgt ctg    240
Ser Gln Ser Gln Asn Asp Leu Met Thr Arg Asp Thr Ala Glu Cys Leu
 65                  70                  75                  80

gcg gag gtg gag ttt gaa gtg aaa ggt gaa gcg tac cgt gca ttc tgg    288
Ala Glu Val Glu Phe Glu Val Lys Gly Glu Ala Tyr Arg Ala Phe Trp
                85                  90                  95

agc cag aat cgg gcg cgt aac caa ccc gac ggt aat ttg cag gtg cca    336
Ser Gln Asn Arg Ala Arg Asn Gln Pro Asp Gly Asn Leu Gln Val Pro
            100                 105                 110

cgc gta gag ctg gcg cgc tgc gcc gac ggc aaa att ctc gcc gac aaa    384
Arg Val Glu Leu Ala Arg Cys Ala Asp Gly Lys Ile Leu Ala Asp Lys
        115                 120                 125

gtg aaa gat aag ctg gaa ctg aca gcg acg tta acc ggg ctg gat tac    432
Val Lys Asp Lys Leu Glu Leu Thr Ala Thr Leu Thr Gly Leu Asp Tyr
    130                 135                 140

ggg cgc ttc acc cgt tcg atg ctg ctt tcg cag ggg caa ttt gct gcc    480
Gly Arg Phe Thr Arg Ser Met Leu Leu Ser Gln Gly Gln Phe Ala Ala
145                 150                 155                 160
```

ttc ctg aat gcc aaa ccc aaa gaa cgc gcg gaa ttg ctc gag gag tta 528
Phe Leu Asn Ala Lys Pro Lys Glu Arg Ala Glu Leu Leu Glu Glu Leu
165 170 175

acc ggc act gaa atc tac ggg caa atc tcg gcg atg gtt ttt gag cag 576
Thr Gly Thr Glu Ile Tyr Gly Gln Ile Ser Ala Met Val Phe Glu Gln
180 185 190

cac aaa tcg gcc cgc aca gag ctg gag aag ctg caa gcg cag gcc agc 624
His Lys Ser Ala Arg Thr Glu Leu Glu Lys Leu Gln Ala Gln Ala Ser
195 200 205

ggc gtc acg ttg ctc acg ccg gaa caa gtg caa tcg ctg aca gcg agt 672
Gly Val Thr Leu Leu Thr Pro Glu Gln Val Gln Ser Leu Thr Ala Ser
210 215 220

ttg cag gta ctt act gac gaa gaa aaa cag tta att acc gcg cag cag 720
Leu Gln Val Leu Thr Asp Glu Glu Lys Gln Leu Ile Thr Ala Gln Gln
225 230 235 240

caa gaa caa caa tcg cta aac tgg tta acg cgt cag gac gaa ttg cag 768
Gln Glu Gln Gln Ser Leu Asn Trp Leu Thr Arg Gln Asp Glu Leu Gln
245 250 255

caa gaa gcc agc cgc cgt cag cag gcc ttg caa cag gcg tta gcc gaa 816
Gln Glu Ala Ser Arg Arg Gln Gln Ala Leu Gln Gln Ala Leu Ala Glu
260 265 270

gaa gaa aaa gcg caa cct caa ctg gcg gcg ctt agt ctg gca caa ccg 864
Glu Glu Lys Ala Gln Pro Gln Leu Ala Ala Leu Ser Leu Ala Gln Pro
275 280 285

gca cga aat ctt cgt cca cac tgg gaa cgc atc gca gaa cac agc gcg 912
Ala Arg Asn Leu Arg Pro His Trp Glu Arg Ile Ala Glu His Ser Ala
290 295 300

gcg ctg gcg cat att cgc cag cag att gaa gaa gta aat act cgc tta 960
Ala Leu Ala His Ile Arg Gln Gln Ile Glu Glu Val Asn Thr Arg Leu
305 310 315 320

cag agc aca atg gcg ctt cgc gcg agc att cgc cac cac gcg gcg aag 1008
Gln Ser Thr Met Ala Leu Arg Ala Ser Ile Arg His His Ala Ala Lys
325 330 335

cag tca gca gaa tta cag cag cag caa caa agc ctg aat acc tgg tta 1056
Gln Ser Ala Glu Leu Gln Gln Gln Gln Gln Ser Leu Asn Thr Trp Leu
340 345 350

cag gaa cac gac cgc ttc cgt cag tgg aac aac gaa ccg gcg ggt tgg 1104
Gln Glu His Asp Arg Phe Arg Gln Trp Asn Asn Glu Pro Ala Gly Trp
355 360 365

cgt gcg cag ttc tcc caa caa acc agc gat cgc gag cat ctg cgg caa 1152
Arg Ala Gln Phe Ser Gln Gln Thr Ser Asp Arg Glu His Leu Arg Gln
370 375 380

tgg cag caa cag tta acc cat gct gag caa aaa ctt aat gcg ctt gcg 1200
Trp Gln Gln Gln Leu Thr His Ala Glu Gln Lys Leu Asn Ala Leu Ala
385 390 395 400

gcg atc acg ttg acg tta acc gcc gat gaa gtt gct acc gcc ctg gcg 1248
Ala Ile Thr Leu Thr Leu Thr Ala Asp Glu Val Ala Thr Ala Leu Ala
405 410 415

caa cat gct gag caa cgc cca ctg cgt cag cac ctg gtc gcg ctg cat 1296
Gln His Ala Glu Gln Arg Pro Leu Arg Gln His Leu Val Ala Leu His
420 425 430

gga cag att gtt ccc caa caa aaa cgt ctg gcg cag tta cag gtc gct 1344
Gly Gln Ile Val Pro Gln Gln Lys Arg Leu Ala Gln Leu Gln Val Ala
435 440 445

atc cag aat gtc acg caa gaa cag acg caa cgt aac gcc gca ctt aac 1392
Ile Gln Asn Val Thr Gln Glu Gln Thr Gln Arg Asn Ala Ala Leu Asn
450 455 460

gaa atg cgc cag cgt tat aaa gaa aag acg cag caa ctt gcc gat gtg 1440
Glu Met Arg Gln Arg Tyr Lys Glu Lys Thr Gln Gln Leu Ala Asp Val
465 470 475 480

aaa acc att tgc gag cag gaa gcg cgc atc aaa acg ctg gaa gct caa 1488
Lys Thr Ile Cys Glu Gln Glu Ala Arg Ile Lys Thr Leu Glu Ala Gln
485 490 495

cgt gca cag tta cag gcg ggt cag cct tgc cca ctt tgt ggt tcc acc 1536
Arg Ala Gln Leu Gln Ala Gly Gln Pro Cys Pro Leu Cys Gly Ser Thr
500 505 510

agc cac ccg gcg gtc gag gcg tat cag gcg ctg gag cct ggc gtt aat 1584
Ser His Pro Ala Val Glu Ala Tyr Gln Ala Leu Glu Pro Gly Val Asn
515 520 525

cag tct cga tta ctg gcg ctg gaa aac gaa gtt aaa aag ctc ggt gaa 1632
Gln Ser Arg Leu Leu Ala Leu Glu Asn Glu Val Lys Lys Leu Gly Glu
530 535 540

gaa ggt gcg acg cta cgt ggg caa ctg gac gcc ata aca aag cag ctt 1680
Glu Gly Ala Thr Leu Arg Gly Gln Leu Asp Ala Ile Thr Lys Gln Leu
545 550 555 560

cag cgt gat gaa aac gaa gcg caa agc ctc cga caa gat gag caa gca 1728
Gln Arg Asp Glu Asn Glu Ala Gln Ser Leu Arg Gln Asp Glu Gln Ala
565 570 575

ctt act caa caa tgg caa gcc gtc acg gcc agc ctc aat atc acc ttg 1776
Leu Thr Gln Gln Trp Gln Ala Val Thr Ala Ser Leu Asn Ile Thr Leu
580 585 590

cag cca ctg gac gat att caa ccg tgg ctg gat gca caa gat gag cac 1824
Gln Pro Leu Asp Asp Ile Gln Pro Trp Leu Asp Ala Gln Asp Glu His
595 600 605

gaa cgc cag ctg cgg tta ctc agc caa cgg cat gaa tta caa ggg cag 1872
Glu Arg Gln Leu Arg Leu Leu Ser Gln Arg His Glu Leu Gln Gly Gln
610 615 620

att gcc gcg cat aat cag caa att atc cag tat caa cag caa att gaa 1920
Ile Ala Ala His Asn Gln Gln Ile Ile Gln Tyr Gln Gln Gln Ile Glu
625 630 635 640

caa cgc cag caa cta ctt tta acg aca ttg acg ggt tat gca ctg aca 1968

Gln Arg Gln Gln Leu Leu Leu Thr Thr Leu Thr Gly Tyr Ala Leu Thr
645 650 655

ttg cca cag gaa gat gaa gaa gag agc tgg ttg gcg aca cgt cag caa 2016
Leu Pro Gln Glu Asp Glu Glu Glu Ser Trp Leu Ala Thr Arg Gln Gln
660 665 670

gaa gcg cag agc tgg cag caa cgc cag aac gaa tta acc gcg ctg caa 2064
Glu Ala Gln Ser Trp Gln Gln Arg Gln Asn Glu Leu Thr Ala Leu Gln
675 680 685

aac cgt att cag cag ctg acg ccg att ctg gaa acg ttg ccg caa agt 2112
Asn Arg Ile Gln Gln Leu Thr Pro Ile Leu Glu Thr Leu Pro Gln Ser
690 695 700

gat gaa ctc ccg cac tgc gaa gaa act gtg gta ttg gaa aac tgg cgg 2160
Asp Glu Leu Pro His Cys Glu Glu Thr Val Val Leu Glu Asn Trp Arg
705 710 715 720

cag gta cat gaa caa tgt ctc gca tta cac agc cag cag cag acg tta 2208
Gln Val His Glu Gln Cys Leu Ala Leu His Ser Gln Gln Gln Thr Leu
725 730 735

cag caa cag gat gtt ctg gcg gcg caa agt ctg caa aaa gcc cag gcg 2256
Gln Gln Gln Asp Val Leu Ala Ala Gln Ser Leu Gln Lys Ala Gln Ala
740 745 750

cag ttt gac acc gcg cta cag gcc agc gtc ttt gac gat cag cag gcg 2304
Gln Phe Asp Thr Ala Leu Gln Ala Ser Val Phe Asp Asp Gln Gln Ala
755 760 765

ttc ctt gcg gcg cta atg gat gaa caa aca cta acg cag ctg gaa cag 2352
Phe Leu Ala Ala Leu Met Asp Glu Gln Thr Leu Thr Gln Leu Glu Gln
770 775 780

ctc aag cag aat ctg gaa aac cag cgc cgt cag gcg caa act ctg gtc 2400
Leu Lys Gln Asn Leu Glu Asn Gln Arg Arg Gln Ala Gln Thr Leu Val
785 790 795 800

act cag aca gca gaa acg ctg gca cag cat caa caa cac cga cct gac 2448
Thr Gln Thr Ala Glu Thr Leu Ala Gln His Gln Gln His Arg Pro Asp
805 810 815

gac ggg ttg gct ctc act gtg acg gtg gag cag att cag caa gag tta 2496
Asp Gly Leu Ala Leu Thr Val Thr Val Glu Gln Ile Gln Gln Glu Leu
820 825 830

gcg caa act cac caa aag ttg cgt gaa aac acc acg agt caa ggc gag 2544
Ala Gln Thr His Gln Lys Leu Arg Glu Asn Thr Thr Ser Gln Gly Glu
835 840 845

att cgc cag cag ctg aag cag gat gca gat aac cgt cag caa caa caa 2592
Ile Arg Gln Gln Leu Lys Gln Asp Ala Asp Asn Arg Gln Gln Gln Gln
850 855 860

acc tta atg cag caa att gct caa atg acg cag cag gtt gag gac tgg 2640
Thr Leu Met Gln Gln Ile Ala Gln Met Thr Gln Gln Val Glu Asp Trp
865 870 875 880

gga tat ctg aat tcg cta ata ggt tcc aaa gag ggc gat aaa ttc cgc 2688
Gly Tyr Leu Asn Ser Leu Ile Gly Ser Lys Glu Gly Asp Lys Phe Arg

885 890 895

aag ttt gcc cag ggg ctg acg ctg gat aat tta gtc cat ctc gct aat 2736
Lys Phe Ala Gln Gly Leu Thr Leu Asp Asn Leu Val His Leu Ala Asn
900 905 910

cag caa ctt acc cgg ctg cac ggg cgc tat ctg tta cag cgc aaa gcc 2784
Gln Gln Leu Thr Arg Leu His Gly Arg Tyr Leu Leu Gln Arg Lys Ala
915 920 925

agc gag gcg ctg gaa gtc gag gtt gtt gat acc tgg cag gca gat gcg 2832
Ser Glu Ala Leu Glu Val Glu Val Val Asp Thr Trp Gln Ala Asp Ala
930 935 940

gta cgc gat acc cgt acc ctt tcc ggc ggc gaa agt ttc ctc gtt agt 2880
Val Arg Asp Thr Arg Thr Leu Ser Gly Gly Glu Ser Phe Leu Val Ser
945 950 955 960

ctg gcg ctg gcg ctg gcg ctt tcg gat ctg gtc agc cat aaa aca cgt 2928
Leu Ala Leu Ala Leu Ala Leu Ser Asp Leu Val Ser His Lys Thr Arg
965 970 975

att gac tcg ctg ttc ctt gat gaa ggt ttt ggc acg ctg gat agc gaa 2976
Ile Asp Ser Leu Phe Leu Asp Glu Gly Phe Gly Thr Leu Asp Ser Glu
980 985 990

acg ctg gat acc gcc ctt gat gcg ctg gat gcc ctg aac gcc agt ggc 3024
Thr Leu Asp Thr Ala Leu Asp Ala Leu Asp Ala Leu Asn Ala Ser Gly
995 1000 1005

aaa acc atc ggt gtg att agc cac gta gaa gcg atg aaa gag cgt att 3072
Lys Thr Ile Gly Val Ile Ser His Val Glu Ala Met Lys Glu Arg Ile
1010 1015 1020

ccg gtg cag atc aaa gtg aaa aag atc aac ggc ctg ggc tac agc aaa 3120
Pro Val Gln Ile Lys Val Lys Lys Ile Asn Gly Leu Gly Tyr Ser Lys
1025 1030 1035 1040

ctg gaa agt acg ttt gca gtg aaa taa 3147
Leu Glu Ser Thr Phe Ala Val Lys *
1045

<210> 217
<211> 1203
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1203)

<400> 217
atg cgc atc ctt cac acc tca gac tgg cat ctc ggc cag aac ttc tac 48
Met Arg Ile Leu His Thr Ser Asp Trp His Leu Gly Gln Asn Phe Tyr
1 5 10 15

agt aaa agc cgc gaa gct gaa cat cag gct ttt ctt gac tgg ctg ctg 96
Ser Lys Ser Arg Glu Ala Glu His Gln Ala Phe Leu Asp Trp Leu Leu
20 25 30

```
gag aca gca caa acc cat cag gtg gat gcg att att gtt gcc ggt gat   144
Glu Thr Ala Gln Thr His Gln Val Asp Ala Ile Ile Val Ala Gly Asp
        35                  40                  45

gtt ttc gat acc ggc tcg ccg ccc agt tac gcc cgc acg tta tac aac   192
Val Phe Asp Thr Gly Ser Pro Pro Ser Tyr Ala Arg Thr Leu Tyr Asn
    50                  55                  60

cgt ttt gtt gtc aat tta cag caa act ggc tgt cat ctg gtg gta ctg   240
Arg Phe Val Val Asn Leu Gln Gln Thr Gly Cys His Leu Val Val Leu
65                  70                  75                  80

gca gga aac cat gac tcg gtc gcc acg ctg aat gaa tcg cgc gat atc   288
Ala Gly Asn His Asp Ser Val Ala Thr Leu Asn Glu Ser Arg Asp Ile
                85                  90                  95

atg gcg ttc ctc aat act acc gtg gtc gcc agc gcc gga cat gcg ccg   336
Met Ala Phe Leu Asn Thr Thr Val Val Ala Ser Ala Gly His Ala Pro
            100                 105                 110

caa atc ttg cct cgt cgc gac ggg acg cca ggc gca gtg ctg tgc ccc   384
Gln Ile Leu Pro Arg Arg Asp Gly Thr Pro Gly Ala Val Leu Cys Pro
        115                 120                 125

att ccg ttt tta cgt ccg cgt gac att att acc agc cag gcg ggg ctt   432
Ile Pro Phe Leu Arg Pro Arg Asp Ile Ile Thr Ser Gln Ala Gly Leu
    130                 135                 140

aac ggt att gaa aaa cag cag cat tta ctg gca gcg att acc gat tat   480
Asn Gly Ile Glu Lys Gln Gln His Leu Leu Ala Ala Ile Thr Asp Tyr
145                 150                 155                 160

tac caa caa cac tat gcc gat gcc tgc aaa ctg cgc ggc gat cag cct   528
Tyr Gln Gln His Tyr Ala Asp Ala Cys Lys Leu Arg Gly Asp Gln Pro
                165                 170                 175

ctg ccc atc atc gcc acg gga cat tta acg acc gtg ggg gcc agt aaa   576
Leu Pro Ile Ile Ala Thr Gly His Leu Thr Thr Val Gly Ala Ser Lys
            180                 185                 190

agt gac gcc gtg cgt gac att tat att ggc acg ctg gac gcg ttt ccg   624
Ser Asp Ala Val Arg Asp Ile Tyr Ile Gly Thr Leu Asp Ala Phe Pro
        195                 200                 205

gca caa aac ttt cca cca gcc gac tac atc gcg ctc ggg cat att cac   672
Ala Gln Asn Phe Pro Pro Ala Asp Tyr Ile Ala Leu Gly His Ile His
    210                 215                 220

cgc gca cag att att ggc ggc atg gaa cat gtt cgc tat tgc ggc tcc   720
Arg Ala Gln Ile Ile Gly Gly Met Glu His Val Arg Tyr Cys Gly Ser
225                 230                 235                 240

ccc att cca ctg agt ttt gat gaa tgc ggt aag agt aaa tat gtc cat   768
Pro Ile Pro Leu Ser Phe Asp Glu Cys Gly Lys Ser Lys Tyr Val His
                245                 250                 255

ctg gtg aca ttt tca aac ggc aaa tta gag agc gtg gaa aac ctg aac   816
Leu Val Thr Phe Ser Asn Gly Lys Leu Glu Ser Val Glu Asn Leu Asn
            260                 265                 270

gta ccg gta acg caa ccc atg gca gtg ctg aaa ggc gat ctg gcg tcg   864
```

Val Pro Val Thr Gln Pro Met Ala Val Leu Lys Gly Asp Leu Ala Ser
275 280 285

att acc gca cag ctg gaa cag tgg cgc gat gta tcg cag gag cca cct 912
Ile Thr Ala Gln Leu Glu Gln Trp Arg Asp Val Ser Gln Glu Pro Pro
290 295 300

gtc tgg ctg gat atc gaa atc act act gat gag tat ctg cat gat att 960
Val Trp Leu Asp Ile Glu Ile Thr Thr Asp Glu Tyr Leu His Asp Ile
305 310 315 320

cag cgc aaa atc cag gca tta acc gaa tca ttg cct gtc gaa gta ttg 1008
Gln Arg Lys Ile Gln Ala Leu Thr Glu Ser Leu Pro Val Glu Val Leu
325 330 335

ctg gta cgt cgg agt cgt gaa cag cgc gag cgt gtg tta gcc agc caa 1056
Leu Val Arg Arg Ser Arg Glu Gln Arg Glu Arg Val Leu Ala Ser Gln
340 345 350

cag cgt gaa acc ctc agc gaa ctc agc gtc gaa gag gtg ttc aat cgc 1104
Gln Arg Glu Thr Leu Ser Glu Leu Ser Val Glu Glu Val Phe Asn Arg
355 360 365

cgt ctg gca ctg gaa gaa ctg gat gaa tcg cag cag caa cgt ctg cag 1152
Arg Leu Ala Leu Glu Glu Leu Asp Glu Ser Gln Gln Gln Arg Leu Gln
370 375 380

cat ctt ttc acc acg acg ttg cat acc ctc gcc gga gaa cac gaa gca 1200
His Leu Phe Thr Thr Thr Leu His Thr Leu Ala Gly Glu His Glu Ala
385 390 395 400

tga 1203
*

<210> 218
<211> 765
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(765)

<400> 218
atg aac gat cta ttt tca ctg gca gga aaa aat atc ttg att acc ggt 48
Met Asn Asp Leu Phe Ser Leu Ala Gly Lys Asn Ile Leu Ile Thr Gly
1 5 10 15

tca gca cag ggc att ggc ttt tta ctg gca acc ggc ctg ggt aaa tat 96
Ser Ala Gln Gly Ile Gly Phe Leu Leu Ala Thr Gly Leu Gly Lys Tyr
20 25 30

ggc gca caa ata att att aat gat att act gcc gaa cgc gca gaa ctt 144
Gly Ala Gln Ile Ile Ile Asn Asp Ile Thr Ala Glu Arg Ala Glu Leu
35 40 45

gct gta gaa aaa ctc cac cag gag ggt att cag gcc gtt gcc gca cct 192
Ala Val Glu Lys Leu His Gln Glu Gly Ile Gln Ala Val Ala Ala Pro
50 55 60

```
ttt aat gtt act cat aaa cat gaa att gat gcc gcc gtt gaa cat atc    240
Phe Asn Val Thr His Lys His Glu Ile Asp Ala Ala Val Glu His Ile
 65                 70                  75                  80

gaa aag gac atc ggc ccc att gat gtg ctg gtg aat aac gcc ggt atc    288
Glu Lys Asp Ile Gly Pro Ile Asp Val Leu Val Asn Asn Ala Gly Ile
                85                  90                  95

cag cgc cgt cat cct ttt act gag ttc cct gaa caa gag tgg aat gat    336
Gln Arg Arg His Pro Phe Thr Glu Phe Pro Glu Gln Glu Trp Asn Asp
            100                 105                 110

gtg atc gca gta aac cag acc gcc gtg ttc ctg gta tcg caa gcg gta    384
Val Ile Ala Val Asn Gln Thr Ala Val Phe Leu Val Ser Gln Ala Val
        115                 120                 125

act cgt cac atg gtt gaa cgc aag gca ggt aaa gtt att aat att tgc    432
Thr Arg His Met Val Glu Arg Lys Ala Gly Lys Val Ile Asn Ile Cys
    130                 135                 140

tcg atg caa agc gaa ctg gga cgt gac acc atc acc cct tat gcc gca    480
Ser Met Gln Ser Glu Leu Gly Arg Asp Thr Ile Thr Pro Tyr Ala Ala
145                 150                 155                 160

tcg aaa ggg gcg gta aaa atg ctc acc cgc ggc atg tgc gtc gag ctg    528
Ser Lys Gly Ala Val Lys Met Leu Thr Arg Gly Met Cys Val Glu Leu
                165                 170                 175

gcg cgc cac aat att cag gtc aac ggt att gcg ccg ggc tat ttc aaa    576
Ala Arg His Asn Ile Gln Val Asn Gly Ile Ala Pro Gly Tyr Phe Lys
            180                 185                 190

aca gaa atg act aaa gca ctg gtt gag gac gaa gcc ttc acc gcc tgg    624
Thr Glu Met Thr Lys Ala Leu Val Glu Asp Glu Ala Phe Thr Ala Trp
        195                 200                 205

ttg tgc aaa cgg acc ccc gcc gca cgc tgg gga gat ccg cag gaa ctg    672
Leu Cys Lys Arg Thr Pro Ala Ala Arg Trp Gly Asp Pro Gln Glu Leu
    210                 215                 220

att ggt gct gcg gtg ttc ctt tct tca aaa gcc tct gat ttc gta aac    720
Ile Gly Ala Ala Val Phe Leu Ser Ser Lys Ala Ser Asp Phe Val Asn
225                 230                 235                 240

ggc cac ctg ttg ttt gtt gat ggc ggc atg tta gtg gct gtt taa        765
Gly His Leu Leu Phe Val Asp Gly Gly Met Leu Val Ala Val  *
                245                 250
```

<210> 219
<211> 1032
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1032)

<400> 219

```
atg caa gtg aaa aca cag tcc tgc gtt gtt gcg ggc aag aaa act gtt    48
Met Gln Val Lys Thr Gln Ser Cys Val Val Ala Gly Lys Lys Thr Val
 1               5                  10                  15

gcc gtt acc gag cag acg ata gat tgg aat aat aat gga aca tta gta    96
Ala Val Thr Glu Gln Thr Ile Asp Trp Asn Asn Asn Gly Thr Leu Val
            20                  25                  30

caa ata acc cga ggt gga att tgc ggt tcc gat tta cat tat tat cag    144
Gln Ile Thr Arg Gly Gly Ile Cys Gly Ser Asp Leu His Tyr Tyr Gln
        35                  40                  45

gaa gga aaa gta ggt aat ttc atg ata aag gca ccg atg gtg tta ggt    192
Glu Gly Lys Val Gly Asn Phe Met Ile Lys Ala Pro Met Val Leu Gly
    50                  55                  60

cat gaa gtt atc ggt aaa gtt att cat agc gac tca tca gaa tta cat    240
His Glu Val Ile Gly Lys Val Ile His Ser Asp Ser Ser Glu Leu His
 65                  70                  75                  80

gaa ggg caa acg gta gcc att aat ccg tct aaa ccg tgc ggt cac tgc    288
Glu Gly Gln Thr Val Ala Ile Asn Pro Ser Lys Pro Cys Gly His Cys
                85                  90                  95

aaa tac tgc att gaa cat aac gag aat cag tgt aca gat atg cgt ttt    336
Lys Tyr Cys Ile Glu His Asn Glu Asn Gln Cys Thr Asp Met Arg Phe
            100                 105                 110

ttt ggc agt gcc atg tat ttc cct cat gtt gat ggt ggt ttt acc cgt    384
Phe Gly Ser Ala Met Tyr Phe Pro His Val Asp Gly Gly Phe Thr Arg
        115                 120                 125

tat aaa atg gtc gaa acg tcg caa tgt gtc cct tat ccg gcc aaa gct    432
Tyr Lys Met Val Glu Thr Ser Gln Cys Val Pro Tyr Pro Ala Lys Ala
    130                 135                 140

gac gaa aag gtt atg gct ttt gcc gaa cct tta gcc gtc gcg att cat    480
Asp Glu Lys Val Met Ala Phe Ala Glu Pro Leu Ala Val Ala Ile His
145                 150                 155                 160

gcc gca cat cag gcc ggc gag tta cag ggc aag cga gta ttt att tcc    528
Ala Ala His Gln Ala Gly Glu Leu Gln Gly Lys Arg Val Phe Ile Ser
                165                 170                 175

ggt gtt gga ccc att ggc tgc ctg att gtc agt gca gtg aaa aca ctg    576
Gly Val Gly Pro Ile Gly Cys Leu Ile Val Ser Ala Val Lys Thr Leu
            180                 185                 190

ggg gcc gcg gaa att gtc tgt gct gat gtg agt ccc cgt tcc ctt tcg    624
Gly Ala Ala Glu Ile Val Cys Ala Asp Val Ser Pro Arg Ser Leu Ser
        195                 200                 205

ctg ggc aaa gag atg ggg gcg gat gtg ctc gta aac cca caa aac gac    672
Leu Gly Lys Glu Met Gly Ala Asp Val Leu Val Asn Pro Gln Asn Asp
    210                 215                 220

gac atg gat cac tgg aaa gcg gaa aaa ggc tat ttc gat gtc agc ttt    720
Asp Met Asp His Trp Lys Ala Glu Lys Gly Tyr Phe Asp Val Ser Phe
225                 230                 235                 240

gaa gtg tcc ggt cat cct tca tca gtg aat acc tgt ctg gag gtc act    768
```

```
Glu Val Ser Gly His Pro Ser Ser Val Asn Thr Cys Leu Glu Val Thr
                245                 250                 255

cgt gca cgc ggc gta atg gtg cag gta ggt atg gga ggc gcg atg gca      816
Arg Ala Arg Gly Val Met Val Gln Val Gly Met Gly Gly Ala Met Ala
            260                 265                 270

gaa ttc cca atg atg acg ttg att ggt aag gag att tca ctc aga ggc      864
Glu Phe Pro Met Met Thr Leu Ile Gly Lys Glu Ile Ser Leu Arg Gly
        275                 280                 285

tct ttc cgt ttt acc agc gaa ttt aat acc gca gtg tca tgg ctg gcg      912
Ser Phe Arg Phe Thr Ser Glu Phe Asn Thr Ala Val Ser Trp Leu Ala
    290                 295                 300

aat ggc gtt atc aat cca ctg cct tta ctg agt gct gaa tat ccc ttc      960
Asn Gly Val Ile Asn Pro Leu Pro Leu Leu Ser Ala Glu Tyr Pro Phe
305                 310                 315                 320

act gac ctg gaa gag gcg cta cgt ttc gcc ggt gat aaa acc cag gca      1008
Thr Asp Leu Glu Glu Ala Leu Arg Phe Ala Gly Asp Lys Thr Gln Ala
                325                 330                 335

gca aaa gtc cag ctt gtt ttc taa                                      1032
Ala Lys Val Gln Leu Val Phe  *
            340


<210> 220
<211> 1647
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1647)

<400> 220
atg gat tcg caa cgc aat ctt tta gtc atc gct ttg ctg ttc gtg tct      48
Met Asp Ser Gln Arg Asn Leu Leu Val Ile Ala Leu Leu Phe Val Ser
 1               5                  10                  15

ttc atg atc tgg caa gcc tgg gag cag gat aaa aac ccg caa cct cag      96
Phe Met Ile Trp Gln Ala Trp Glu Gln Asp Lys Asn Pro Gln Pro Gln
            20                  25                  30

gcc caa cag acc acg cag aca acg acc acc gca gcg ggt agc gcc gcc      144
Ala Gln Gln Thr Thr Gln Thr Thr Thr Thr Ala Ala Gly Ser Ala Ala
        35                  40                  45

gac cag ggc gta ccg gcc agt ggc cag ggg aaa ctg atc tcg gtt aag      192
Asp Gln Gly Val Pro Ala Ser Gly Gln Gly Lys Leu Ile Ser Val Lys
    50                  55                  60

acc gac gtg ctt gat ctg acc atc aac acc cgt ggt ggt gat gtt gag      240
Thr Asp Val Leu Asp Leu Thr Ile Asn Thr Arg Gly Gly Asp Val Glu
65                  70                  75                  80

caa gct ctg ctg cct gct tac ccg aaa gag ctg aac tct acc cag ccg      288
Gln Ala Leu Leu Pro Ala Tyr Pro Lys Glu Leu Asn Ser Thr Gln Pro
                85                  90                  95
```

```
ttc cag ctg ttg gaa act tca ccg cag ttt att tat cag gca cag agc    336
Phe Gln Leu Leu Glu Thr Ser Pro Gln Phe Ile Tyr Gln Ala Gln Ser
            100                 105                 110

ggt ctg acc ggt cgt gat ggc ccg gat aac ccg gct aac ggc ccg cgt    384
Gly Leu Thr Gly Arg Asp Gly Pro Asp Asn Pro Ala Asn Gly Pro Arg
        115                 120                 125

ccg ctg tat aac gtt gaa aaa gac gct tat gtg ctg gct gaa ggt caa    432
Pro Leu Tyr Asn Val Glu Lys Asp Ala Tyr Val Leu Ala Glu Gly Gln
    130                 135                 140

aac gaa ctg cag gtg ccg atg acg tat acc gac gcg gca ggc aac acg    480
Asn Glu Leu Gln Val Pro Met Thr Tyr Thr Asp Ala Ala Gly Asn Thr
145                 150                 155                 160

ttt acc aaa acg ttt gtc ctg aaa cgt ggt gat tac gct gtc aac gtc    528
Phe Thr Lys Thr Phe Val Leu Lys Arg Gly Asp Tyr Ala Val Asn Val
                165                 170                 175

aac tac aac gtg cag aac gct ggc gag aaa ccg ctg gaa atc tcc tcg    576
Asn Tyr Asn Val Gln Asn Ala Gly Glu Lys Pro Leu Glu Ile Ser Ser
            180                 185                 190

ttt ggt cag ttg aag caa tcc atc act ctg cca ccg cat ctc gat acc    624
Phe Gly Gln Leu Lys Gln Ser Ile Thr Leu Pro Pro His Leu Asp Thr
        195                 200                 205

gga agc agc aac ttc gca ctg cac acc ttc cgt ggc gcg gcg tac tcc    672
Gly Ser Ser Asn Phe Ala Leu His Thr Phe Arg Gly Ala Ala Tyr Ser
    210                 215                 220

acg cct gac gag aag tat gag aaa tac aag ttc gat acc att gcc gat    720
Thr Pro Asp Glu Lys Tyr Glu Lys Tyr Lys Phe Asp Thr Ile Ala Asp
225                 230                 235                 240

aac gaa aac ctg aac atc tct tcg aaa ggt ggt tgg gtg gcg atg ctg    768
Asn Glu Asn Leu Asn Ile Ser Ser Lys Gly Gly Trp Val Ala Met Leu
                245                 250                 255

caa cag tat ttc gcg acg gcg tgg atc ccg cat aac gac ggt acc aac    816
Gln Gln Tyr Phe Ala Thr Ala Trp Ile Pro His Asn Asp Gly Thr Asn
            260                 265                 270

aac ttc tat acc gct aat ctg ggt aac ggc atc gcc gct atc ggc tat    864
Asn Phe Tyr Thr Ala Asn Leu Gly Asn Gly Ile Ala Ala Ile Gly Tyr
        275                 280                 285

aaa tct cag ccg gta ctg gtt cag cct ggt cag act ggc gcg atg aac    912
Lys Ser Gln Pro Val Leu Val Gln Pro Gly Gln Thr Gly Ala Met Asn
    290                 295                 300

agc acc ctg tgg gtt ggc ccg gaa atc cag gac aaa atg gca gct gtt    960
Ser Thr Leu Trp Val Gly Pro Glu Ile Gln Asp Lys Met Ala Ala Val
305                 310                 315                 320

gct ccg cac ctg gat ctg acc gtt gat tac ggt tgg ttg tgg ttc atc    1008
Ala Pro His Leu Asp Leu Thr Val Asp Tyr Gly Trp Leu Trp Phe Ile
                325                 330                 335
```

tct cag ccg ctg ttc aaa ctg ctg aaa tgg atc cat agc ttt gtg ggt 1056
Ser Gln Pro Leu Phe Lys Leu Leu Lys Trp Ile His Ser Phe Val Gly
340 345 350

aac tgg ggc ttc tcc att atc atc atc acc ttt atc gtt cgt ggc atc 1104
Asn Trp Gly Phe Ser Ile Ile Ile Ile Thr Phe Ile Val Arg Gly Ile
355 360 365

atg tac ccg ctg acc aaa gcg cag tac acc tcc atg gcg aag atg cgt 1152
Met Tyr Pro Leu Thr Lys Ala Gln Tyr Thr Ser Met Ala Lys Met Arg
370 375 380

atg ttg cag ccg aag att cag gca atg cgt gag cgt ctg ggc gat gac 1200
Met Leu Gln Pro Lys Ile Gln Ala Met Arg Glu Arg Leu Gly Asp Asp
385 390 395 400

aaa cag cgt atc agc cag gaa atg atg gcg ctg tac aaa gct gag aag 1248
Lys Gln Arg Ile Ser Gln Glu Met Met Ala Leu Tyr Lys Ala Glu Lys
405 410 415

gtt aac ccg ctg ggc ggc tgc ttc ccg ctg ctg atc cag atg cca atc 1296
Val Asn Pro Leu Gly Gly Cys Phe Pro Leu Leu Ile Gln Met Pro Ile
420 425 430

ttc ctg gcg ttg tac tac atg ctg atg ggt tcc gtt gaa ctg cgt cag 1344
Phe Leu Ala Leu Tyr Tyr Met Leu Met Gly Ser Val Glu Leu Arg Gln
435 440 445

gca ccg ttt gca ctg tgg atc cac gac ctg tcg gca cag gac ccg tac 1392
Ala Pro Phe Ala Leu Trp Ile His Asp Leu Ser Ala Gln Asp Pro Tyr
450 455 460

tac atc ctg ccg atc ctg atg ggc gta acg atg ttc ttc att cag aag 1440
Tyr Ile Leu Pro Ile Leu Met Gly Val Thr Met Phe Phe Ile Gln Lys
465 470 475 480

atg tcg ccg acc aca gtg acc gac ccg atg cag cag aag atc atg acc 1488
Met Ser Pro Thr Thr Val Thr Asp Pro Met Gln Gln Lys Ile Met Thr
485 490 495

ttt atg ccg gtc atc ttc acc gtg ttc ttc ctg tgg ttc ccg tca ggt 1536
Phe Met Pro Val Ile Phe Thr Val Phe Phe Leu Trp Phe Pro Ser Gly
500 505 510

ctg gtg ctg tac tat atc gtc agc aac ctg gta acc att att cag cag 1584
Leu Val Leu Tyr Tyr Ile Val Ser Asn Leu Val Thr Ile Ile Gln Gln
515 520 525

cag ctg att tac cgt ggt ctg gaa aaa cgt ggc ctg cat agc cgc gag 1632
Gln Leu Ile Tyr Arg Gly Leu Glu Lys Arg Gly Leu His Ser Arg Glu
530 535 540

aag aaa aaa tcc tga 1647
Lys Lys Lys Ser *
545

<210> 221
<211> 1629
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1629)

<400> 221
atg aaa act cgc gac tcg caa tca agt gac gtg att atc att ggc ggc 48
Met Lys Thr Arg Asp Ser Gln Ser Ser Asp Val Ile Ile Ile Gly Gly
1 5 10 15

ggc gca acg gga gcc ggg att gcc cgc gac tgt gcc ctg cgc ggg ctg 96
Gly Ala Thr Gly Ala Gly Ile Ala Arg Asp Cys Ala Leu Arg Gly Leu
20 25 30

cgc gtg att ttg gtt gag cgc cac gac atc gca acc ggt gcc acc ggg 144
Arg Val Ile Leu Val Glu Arg His Asp Ile Ala Thr Gly Ala Thr Gly
35 40 45

cgt aac cac ggc ctg ctg cac agc ggt gcg cgc tat gcg gta acc gat 192
Arg Asn His Gly Leu Leu His Ser Gly Ala Arg Tyr Ala Val Thr Asp
50 55 60

gcg gaa tcg gcc cgc gaa tgc att agt gaa aac cag atc ctg aaa cgc 240
Ala Glu Ser Ala Arg Glu Cys Ile Ser Glu Asn Gln Ile Leu Lys Arg
65 70 75 80

att gca cgt cac tgc gtt gaa cca acc aac ggc ctg ttt atc acc ctg 288
Ile Ala Arg His Cys Val Glu Pro Thr Asn Gly Leu Phe Ile Thr Leu
85 90 95

ccg gaa gat gac ctc tcc ttc cag gcc act ttt att cgc gcc tgc gaa 336
Pro Glu Asp Asp Leu Ser Phe Gln Ala Thr Phe Ile Arg Ala Cys Glu
100 105 110

gaa gca ggg atc agc gca gaa gct ata gac ccg cag caa gcg cgc att 384
Glu Ala Gly Ile Ser Ala Glu Ala Ile Asp Pro Gln Gln Ala Arg Ile
115 120 125

atc gaa cct gcc gtt aac ccg gca ctg att ggc gcg gtg aaa gtt ccg 432
Ile Glu Pro Ala Val Asn Pro Ala Leu Ile Gly Ala Val Lys Val Pro
130 135 140

gat ggc acc gtt gat cca ttt cgt ctg acc gca gca aac atg ctg gat 480
Asp Gly Thr Val Asp Pro Phe Arg Leu Thr Ala Ala Asn Met Leu Asp
145 150 155 160

gcc aaa gaa cac ggt gcc gtt atc ctt acc gct cat gaa gtc acg ggg 528
Ala Lys Glu His Gly Ala Val Ile Leu Thr Ala His Glu Val Thr Gly
165 170 175

ctg att cgt gaa ggc gcg acg gtg tgc ggt gtt cgt gta cgt aac cat 576
Leu Ile Arg Glu Gly Ala Thr Val Cys Gly Val Arg Val Arg Asn His
180 185 190

ctc acc ggc gaa act cag gcc ctt cat gca cct gtc gtg gtt aat gcc 624
Leu Thr Gly Glu Thr Gln Ala Leu His Ala Pro Val Val Val Asn Ala
195 200 205

gct ggg atc tgg ggg caa cac att gcc gaa tat gcc gat ctg cgc att 672
Ala Gly Ile Trp Gly Gln His Ile Ala Glu Tyr Ala Asp Leu Arg Ile
210 215 220

```
cgc atg ttc ccg gcg aaa gga tcg ctg ctg atc atg gat cac cgc att    720
Arg Met Phe Pro Ala Lys Gly Ser Leu Leu Ile Met Asp His Arg Ile
225                 230                 235                 240

aac cag cat gtg atc aac cgc tgc cgt aaa cct tcc gac gcc gat att    768
Asn Gln His Val Ile Asn Arg Cys Arg Lys Pro Ser Asp Ala Asp Ile
                245                 250                 255

ctg gtg cct ggc gat acc att tcg ctg att ggt acc acc tct tta cgt    816
Leu Val Pro Gly Asp Thr Ile Ser Leu Ile Gly Thr Thr Ser Leu Arg
            260                 265                 270

att gat tac aac gag att gac gat aat cga gtg acg gca gaa gag gtt    864
Ile Asp Tyr Asn Glu Ile Asp Asp Asn Arg Val Thr Ala Glu Glu Val
        275                 280                 285

gat att ctg ctg cgt gaa ggg gaa aaa ctg gcc ccc gtg atg gcg aaa    912
Asp Ile Leu Leu Arg Glu Gly Glu Lys Leu Ala Pro Val Met Ala Lys
    290                 295                 300

acg cgc att ttg cgg gcc tat tct ggc gtg cgc ccg ctg gtt gcc agc    960
Thr Arg Ile Leu Arg Ala Tyr Ser Gly Val Arg Pro Leu Val Ala Ser
305                 310                 315                 320

gat gac gac ccg agc gga cgt aac gtc agc cgt ggc atc gtg ctg ctc    1008
Asp Asp Asp Pro Ser Gly Arg Asn Val Ser Arg Gly Ile Val Leu Leu
                325                 330                 335

gac cat gct gaa cgc gat ggt ctg gac gga ttt atc acc atc acc ggt    1056
Asp His Ala Glu Arg Asp Gly Leu Asp Gly Phe Ile Thr Ile Thr Gly
            340                 345                 350

ggc aaa ctg atg acc tat cgg ctg atg gct gaa tgg gct acc gac gcg    1104
Gly Lys Leu Met Thr Tyr Arg Leu Met Ala Glu Trp Ala Thr Asp Ala
        355                 360                 365

gta tgc cgc aaa ctg ggc aac acg cgc ccc tgt acg act gcc gat ctg    1152
Val Cys Arg Lys Leu Gly Asn Thr Arg Pro Cys Thr Thr Ala Asp Leu
    370                 375                 380

gca ctg cct ggt tca caa gaa ccc gct gaa gtt acc ttg cgt aaa gtc    1200
Ala Leu Pro Gly Ser Gln Glu Pro Ala Glu Val Thr Leu Arg Lys Val
385                 390                 395                 400

atc tcc ctg cct gcc ccg ctg cgc ggt tct gcg gtt tat cgt cat ggc    1248
Ile Ser Leu Pro Ala Pro Leu Arg Gly Ser Ala Val Tyr Arg His Gly
                405                 410                 415

gat cgc acg cct gcc tgg ctg agc gaa ggc cgt ctg cac cgt agc ctg    1296
Asp Arg Thr Pro Ala Trp Leu Ser Glu Gly Arg Leu His Arg Ser Leu
            420                 425                 430

gta tgt gag tgc gaa gcg gta act gcg ggt gaa gtg cag tac gcg gta    1344
Val Cys Glu Cys Glu Ala Val Thr Ala Gly Glu Val Gln Tyr Ala Val
        435                 440                 445

gaa aat tta aac gtt aat agc ctg ctg gat tta cgc cgt cgt acc cgt    1392
Glu Asn Leu Asn Val Asn Ser Leu Leu Asp Leu Arg Arg Arg Thr Arg
    450                 455                 460
```

```
gtg ggg atg ggc acc tgc cag ggc gaa ctc tgc gcc tgc cgc gct gcc    1440
Val Gly Met Gly Thr Cys Gln Gly Glu Leu Cys Ala Cys Arg Ala Ala
465                 470                 475                 480

gga ctg ctg caa cgt ttt aac gtc acg acg tcc gcg caa tct atc gag    1488
Gly Leu Leu Gln Arg Phe Asn Val Thr Thr Ser Ala Gln Ser Ile Glu
                485                 490                 495

caa ctt tcc acc ttc ctt aac gaa cgc tgg aaa ggc gtg caa ccc atc    1536
Gln Leu Ser Thr Phe Leu Asn Glu Arg Trp Lys Gly Val Gln Pro Ile
            500                 505                 510

gcc tgg gga gat gca ctg cgc gaa agc gaa ttt acc cgc tgg gtt tat    1584
Ala Trp Gly Asp Ala Leu Arg Glu Ser Glu Phe Thr Arg Trp Val Tyr
        515                 520                 525

cag gga ttg tgt ggt ctg gag aag gag cag aaa gat gcg ctt tga        1629
Gln Gly Leu Cys Gly Leu Glu Lys Glu Gln Lys Asp Ala Leu  *
    530                 535                 540
```

<210> 222
<211> 1260
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1260)

<400> 222

```
atg cgc ttt gat act gtc att atg ggc ggc ggc ctc gcc gga tta ctc    48
Met Arg Phe Asp Thr Val Ile Met Gly Gly Gly Leu Ala Gly Leu Leu
 1               5                  10                  15

tgt ggc ctg caa ctg caa aaa cac ggc ctg cgc tgt gcc att gtc act    96
Cys Gly Leu Gln Leu Gln Lys His Gly Leu Arg Cys Ala Ile Val Thr
            20                  25                  30

cgt ggt caa agc gca ctg cat ttc tca tcc gga tcg ctg gat ttg ctg    144
Arg Gly Gln Ser Ala Leu His Phe Ser Ser Gly Ser Leu Asp Leu Leu
        35                  40                  45

agc cat ctg cca gat ggt caa ccg gtg aca gac att cac agt gga ctg    192
Ser His Leu Pro Asp Gly Gln Pro Val Thr Asp Ile His Ser Gly Leu
    50                  55                  60

gaa tct ttg cgt cag cag gca cca gcc cat cct tac tcc ctt ctc gag    240
Glu Ser Leu Arg Gln Gln Ala Pro Ala His Pro Tyr Ser Leu Leu Glu
65                  70                  75                  80

cca caa cgc gtg ctc gat ctc gct tgc cag gcg cag gca tta atc gct    288
Pro Gln Arg Val Leu Asp Leu Ala Cys Gln Ala Gln Ala Leu Ile Ala
                85                  90                  95

gaa agc ggt gcg caa ttg cag ggc agc gta gaa ctt gct cac cag cgg    336
Glu Ser Gly Ala Gln Leu Gln Gly Ser Val Glu Leu Ala His Gln Arg
            100                 105                 110

gtt acg ccg ctc ggc act ctg cgc tct acc tgg cta agt tcg cca gaa    384
```

```
Val Thr Pro Leu Gly Thr Leu Arg Ser Thr Trp Leu Ser Ser Pro Glu
        115                 120                 125

gtc ccc gtc tgg ccg ctg ccc gcg aag aaa ata tgt gta gtg gga att      432
Val Pro Val Trp Pro Leu Pro Ala Lys Lys Ile Cys Val Val Gly Ile
    130                 135                 140

agc ggc ctg atg gat ttt cag gcg cac ctt gcg gca gct tcg ttg cgt      480
Ser Gly Leu Met Asp Phe Gln Ala His Leu Ala Ala Ala Ser Leu Arg
145                 150                 155                 160

gaa ctc ggc ctt gcc gtt gaa acc gca gaa ata gag ctg ccg gaa ctg      528
Glu Leu Gly Leu Ala Val Glu Thr Ala Glu Ile Glu Leu Pro Glu Leu
                165                 170                 175

gat gtg ctg cgc aat aac gcc acc gaa ttt cgc gcg gtg aat atc gcc      576
Asp Val Leu Arg Asn Asn Ala Thr Glu Phe Arg Ala Val Asn Ile Ala
            180                 185                 190

cgt ttc ctt gat aat gaa gaa aac tgg ccg ctg tta ctt gat gcg ctt      624
Arg Phe Leu Asp Asn Glu Glu Asn Trp Pro Leu Leu Leu Asp Ala Leu
        195                 200                 205

att cct gtc gcc aat acc tgc gaa atg atc ctg atg ccc gcc tgc ttc      672
Ile Pro Val Ala Asn Thr Cys Glu Met Ile Leu Met Pro Ala Cys Phe
    210                 215                 220

ggt ctg gcc gat gac aaa ctg tgg cgt tgg ttg aat gaa aaa cta cct      720
Gly Leu Ala Asp Asp Lys Leu Trp Arg Trp Leu Asn Glu Lys Leu Pro
225                 230                 235                 240

tgt tca ctg atg ctt ttg cca acg ctg ccg cct tcc gtg ctg ggc att      768
Cys Ser Leu Met Leu Leu Pro Thr Leu Pro Pro Ser Val Leu Gly Ile
                245                 250                 255

cgt ctg caa aac cag tta cag cgc cag ttt gtg cgc cag ggt ggc gtg      816
Arg Leu Gln Asn Gln Leu Gln Arg Gln Phe Val Arg Gln Gly Gly Val
            260                 265                 270

tgg atg ccg ggc gat gaa gtg aaa aaa gtg acc tgt aaa aat ggc gta      864
Trp Met Pro Gly Asp Glu Val Lys Lys Val Thr Cys Lys Asn Gly Val
        275                 280                 285

gtg aac gaa atc tgg acc cgc aat cac gcc gat att ccg cta cgt cca      912
Val Asn Glu Ile Trp Thr Arg Asn His Ala Asp Ile Pro Leu Arg Pro
    290                 295                 300

cgt ttc gcg gtt ctc gcc agc ggc agt ttc ttt agt ggc gga ctg gta      960
Arg Phe Ala Val Leu Ala Ser Gly Ser Phe Phe Ser Gly Gly Leu Val
305                 310                 315                 320

gcg gaa cgt aac ggc att cga gag ccg att ctc ggc ctt gat gtg cta     1008
Ala Glu Arg Asn Gly Ile Arg Glu Pro Ile Leu Gly Leu Asp Val Leu
                325                 330                 335

caa acc gcc acg cgg ggt gaa tgg tat aag gga gat ttt ttt gcg ccg     1056
Gln Thr Ala Thr Arg Gly Glu Trp Tyr Lys Gly Asp Phe Phe Ala Pro
            340                 345                 350

caa ccg tgg cag cag ttc ggt gta acc act gat gag acg cta cgc ccg     1104
Gln Pro Trp Gln Gln Phe Gly Val Thr Thr Asp Glu Thr Leu Arg Pro
```

```
        355                 360                 365

tca cag gca ggg caa acc att gaa aac ctg ttt gcc atc ggt tcg gtg    1152
Ser Gln Ala Gly Gln Thr Ile Glu Asn Leu Phe Ala Ile Gly Ser Val
    370                 375                 380

ctg ggc gga ttt gat ccc atc gcc cag gga tgc ggc ggc ggt gtt tgt    1200
Leu Gly Gly Phe Asp Pro Ile Ala Gln Gly Cys Gly Gly Gly Val Cys
385                 390                 395                 400

gcc gtc agt gct tta cat gcc gct caa cag att gcc caa cgc gca gga    1248
Ala Val Ser Ala Leu His Ala Ala Gln Gln Ile Ala Gln Arg Ala Gly
                405                 410                 415

ggc caa caa tga                                                    1260
Gly Gln Gln  *
```

```
<210> 223
<211> 1191
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1191)

<400> 223
atg aat gac acc agc ttc gaa aac tgc att aag tgc acc gtc tgc acc    48
Met Asn Asp Thr Ser Phe Glu Asn Cys Ile Lys Cys Thr Val Cys Thr
 1               5                  10                  15

acc gcc tgc ccg gtg agc cgg gtg aat ccc ggt tat cca ggg cca aaa    96
Thr Ala Cys Pro Val Ser Arg Val Asn Pro Gly Tyr Pro Gly Pro Lys
            20                  25                  30

caa gcc ggg ccg gat ggc gag cgt ctg cgt ttg aaa gat ggc gca ctg    144
Gln Ala Gly Pro Asp Gly Glu Arg Leu Arg Leu Lys Asp Gly Ala Leu
        35                  40                  45

tat gac gag gcg ctg aaa tat tgc atc aac tgc aaa cgt tgt gaa gtc    192
Tyr Asp Glu Ala Leu Lys Tyr Cys Ile Asn Cys Lys Arg Cys Glu Val
    50                  55                  60

gcc tgc ccg tcc gat gtg aag att ggc gat att atc cag cgc gcg cgg    240
Ala Cys Pro Ser Asp Val Lys Ile Gly Asp Ile Ile Gln Arg Ala Arg
65                  70                  75                  80

gcg aaa tat gac acc acg cgc ccg tcg ctg cgt aat ttt gtg ttg agt    288
Ala Lys Tyr Asp Thr Thr Arg Pro Ser Leu Arg Asn Phe Val Leu Ser
                85                  90                  95

cat acc gac ctg atg ggt agc gtt tcc acg ccg ttc gca cca atc gtc    336
His Thr Asp Leu Met Gly Ser Val Ser Thr Pro Phe Ala Pro Ile Val
            100                 105                 110

aac acc gct acc tcg ctg aaa ccg gtg cgg cag ctg ctt gat gcg gcg    384
Asn Thr Ala Thr Ser Leu Lys Pro Val Arg Gln Leu Leu Asp Ala Ala
        115                 120                 125
```

tta aaa atc gat cat cgc cgc acg cta ccg aaa tac tcc ttc ggc acg 432
Leu Lys Ile Asp His Arg Arg Thr Leu Pro Lys Tyr Ser Phe Gly Thr
130 135 140

ttc cgt cgc tgg tat cgc agc gtg gcg gct cag caa gca caa tat aaa 480
Phe Arg Arg Trp Tyr Arg Ser Val Ala Ala Gln Gln Ala Gln Tyr Lys
145 150 155 160

gac cag gtc gct ttc ttt cac ggc tgc ttc gtt aac tac aac cat ccg 528
Asp Gln Val Ala Phe Phe His Gly Cys Phe Val Asn Tyr Asn His Pro
165 170 175

cag tta ggt aaa gat tta att aaa gtg ctc aac gca atg ggt acc ggt 576
Gln Leu Gly Lys Asp Leu Ile Lys Val Leu Asn Ala Met Gly Thr Gly
180 185 190

gta caa ctg ctc agc aaa gaa aaa tgc tgc ggc gta ccg cta atc gcc 624
Val Gln Leu Leu Ser Lys Glu Lys Cys Cys Gly Val Pro Leu Ile Ala
195 200 205

aac ggc ttt acc gat aaa gca cgc aaa cag gca att acg aat gta gag 672
Asn Gly Phe Thr Asp Lys Ala Arg Lys Gln Ala Ile Thr Asn Val Glu
210 215 220

tcg atc cgc gaa gct gtg gga gta aaa ggc att ccg gtg att gcc acc 720
Ser Ile Arg Glu Ala Val Gly Val Lys Gly Ile Pro Val Ile Ala Thr
225 230 235 240

tcc tca acc tgt aca ttt gcc ctg cgc gac gaa tac ccg gaa gtg ctg 768
Ser Ser Thr Cys Thr Phe Ala Leu Arg Asp Glu Tyr Pro Glu Val Leu
245 250 255

aat gtc gac aac aaa ggc ttg cgc gat cat atc gaa ctg gca acc cgc 816
Asn Val Asp Asn Lys Gly Leu Arg Asp His Ile Glu Leu Ala Thr Arg
260 265 270

tgg ctg tgg cgc aag ctg gac gaa ggc aaa acg tta ccg ctg aaa ccg 864
Trp Leu Trp Arg Lys Leu Asp Glu Gly Lys Thr Leu Pro Leu Lys Pro
275 280 285

ctg ccg ctg aaa gtg gtt tat cac act ccg tgc cat atg gaa aaa atg 912
Leu Pro Leu Lys Val Val Tyr His Thr Pro Cys His Met Glu Lys Met
290 295 300

ggc tgg acg ctc tac acc ctg gag ctg ttg cgt aac atc ccg ggg ctt 960
Gly Trp Thr Leu Tyr Thr Leu Glu Leu Leu Arg Asn Ile Pro Gly Leu
305 310 315 320

gag tta acg gtg ctg gat tcc cag tgc tgc ggt att gcg ggt act tac 1008
Glu Leu Thr Val Leu Asp Ser Gln Cys Cys Gly Ile Ala Gly Thr Tyr
325 330 335

ggt ttc aaa aaa gag aac tac ccc acc tca caa gcc atc ggc gca cca 1056
Gly Phe Lys Lys Glu Asn Tyr Pro Thr Ser Gln Ala Ile Gly Ala Pro
340 345 350

ctg ttc cgc cag ata gaa gaa agc ggc gca gat ctg gtg gtc acc gac 1104
Leu Phe Arg Gln Ile Glu Glu Ser Gly Ala Asp Leu Val Val Thr Asp
355 360 365

tgc gaa acc tgt aaa tgg cag att gag atg tcc aca agt ctt cgc tgc 1152

```
Cys Glu Thr Cys Lys Trp Gln Ile Glu Met Ser Thr Ser Leu Arg Cys
    370                 375                 380

gaa cat ccg att acg cta ctg gcc cag gcg ctg gct taa                 1191
Glu His Pro Ile Thr Leu Leu Ala Gln Ala Leu Ala  *
385                 390                 395


<210> 224
<211> 909
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(909)

<400> 224
atg caa tta cgt aaa cct gcc aca gca atc ctc gcc ctg gcg ctt tcc      48
Met Gln Leu Arg Lys Pro Ala Thr Ala Ile Leu Ala Leu Ala Leu Ser
 1               5                  10                  15

gca gga ctg gca cag gca gat gac gcc gcc ccg gca gcg ggc agt act      96
Ala Gly Leu Ala Gln Ala Asp Asp Ala Ala Pro Ala Ala Gly Ser Thr
             20                  25                  30

ctg gac aaa atc gcc aaa aac ggt gtg att gtc gtc ggt cac cgt gaa     144
Leu Asp Lys Ile Ala Lys Asn Gly Val Ile Val Val Gly His Arg Glu
         35                  40                  45

tct tca gtg cct ttc tct tat tac gac aat cag caa aaa gtg gtg ggt     192
Ser Ser Val Pro Phe Ser Tyr Tyr Asp Asn Gln Gln Lys Val Val Gly
     50                  55                  60

tac tcg cag gat tac tcc aac gcc att gtt gaa gca gtg aaa aag aaa     240
Tyr Ser Gln Asp Tyr Ser Asn Ala Ile Val Glu Ala Val Lys Lys Lys
 65                  70                  75                  80

ctc aac aaa ccg gac ttg cag gta aaa ctg att ccg att acc tca caa     288
Leu Asn Lys Pro Asp Leu Gln Val Lys Leu Ile Pro Ile Thr Ser Gln
                 85                  90                  95

aac cgt att cca ctg ctg caa aac ggc act ttc gat ttt gaa tgt ggt     336
Asn Arg Ile Pro Leu Leu Gln Asn Gly Thr Phe Asp Phe Glu Cys Gly
            100                 105                 110

tct acc acc aac aac gtc gaa cgc caa aaa cag gcg gct ttc tct gac     384
Ser Thr Thr Asn Asn Val Glu Arg Gln Lys Gln Ala Ala Phe Ser Asp
        115                 120                 125

act att ttc gtg gtc ggt acg cgc ctg ttg acc aaa aag ggt ggc gat     432
Thr Ile Phe Val Val Gly Thr Arg Leu Leu Thr Lys Lys Gly Gly Asp
    130                 135                 140

atc aaa gat ttt gcc aac ctg aaa gac aaa gcc gta gtc gtc act tcc     480
Ile Lys Asp Phe Ala Asn Leu Lys Asp Lys Ala Val Val Val Thr Ser
145                 150                 155                 160

ggc act acc tct gaa gtt ttg ctc aac aaa ctg aat gaa gag caa aaa     528
Gly Thr Thr Ser Glu Val Leu Leu Asn Lys Leu Asn Glu Glu Gln Lys
                165                 170                 175
```

```
atg aat atg cgc atc atc agc gcc aaa gat cac ggt gac tct ttc cgc     576
Met Asn Met Arg Ile Ile Ser Ala Lys Asp His Gly Asp Ser Phe Arg
            180                 185                 190

acc ctg gaa agc ggt cgt gcc gtt gcc ttt atg atg gat gac gct ctg     624
Thr Leu Glu Ser Gly Arg Ala Val Ala Phe Met Met Asp Asp Ala Leu
        195                 200                 205

ctg gcc ggt gaa cgt gcg aaa gcg aag aaa cca gac aac tgg gaa atc     672
Leu Ala Gly Glu Arg Ala Lys Ala Lys Lys Pro Asp Asn Trp Glu Ile
    210                 215                 220

gtc ggc aag ccg cag tct cag gag gcc tac ggt tgt atg ttg cgt aaa     720
Val Gly Lys Pro Gln Ser Gln Glu Ala Tyr Gly Cys Met Leu Arg Lys
225                 230                 235                 240

gat gat ccg cag ttc aaa aag ctg atg gat gac acc atc gct cag gtg     768
Asp Asp Pro Gln Phe Lys Lys Leu Met Asp Asp Thr Ile Ala Gln Val
                245                 250                 255

cag acc tcc ggt gaa gcg gaa aaa tgg ttt gat aag tgg ttc aaa aat     816
Gln Thr Ser Gly Glu Ala Glu Lys Trp Phe Asp Lys Trp Phe Lys Asn
            260                 265                 270

cca att ccg ccg aaa aac ctg aac atg aat ttc gaa ctg tca gac gaa     864
Pro Ile Pro Pro Lys Asn Leu Asn Met Asn Phe Glu Leu Ser Asp Glu
        275                 280                 285

atg aaa gca ctg ttc aaa gaa ccg aat gac aag gca ctg aac taa         909
Met Lys Ala Leu Phe Lys Glu Pro Asn Asp Lys Ala Leu Asn  *
    290                 295                 300
```

```
<210> 225
<211> 987
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(987)

<400> 225
atg aat aac agc gct ttt act ttc cag aca cta cac ccg gat acc atc     48
Met Asn Asn Ser Ala Phe Thr Phe Gln Thr Leu His Pro Asp Thr Ile
 1               5                  10                  15

atg gac gct ctg ttt gag cat ggg atc cgg gtg gat tcc ggt ctt acc     96
Met Asp Ala Leu Phe Glu His Gly Ile Arg Val Asp Ser Gly Leu Thr
            20                  25                  30

ccg ctt aac agc tat gaa aac cgt gtc tat caa ttt cag gac gaa gat     144
Pro Leu Asn Ser Tyr Glu Asn Arg Val Tyr Gln Phe Gln Asp Glu Asp
        35                  40                  45

cgt cga cgt ttt gtc gtc aaa ttt tat cgc cct gaa cgt tgg aca gcc     192
Arg Arg Arg Phe Val Val Lys Phe Tyr Arg Pro Glu Arg Trp Thr Ala
    50                  55                  60
```

```
gat caa atc ctc gaa gaa cat caa ttt gcg ttg cag ctg gta aat gat      240
Asp Gln Ile Leu Glu Glu His Gln Phe Ala Leu Gln Leu Val Asn Asp
 65                  70                  75                  80

gaa gtt ccg gtc gca gca cct gtg gcc ttt aac ggt cag act tta ttg      288
Glu Val Pro Val Ala Ala Pro Val Ala Phe Asn Gly Gln Thr Leu Leu
                 85                  90                  95

aat cat cag gga ttt tat ttc gct gtt ttt cca agc gtc ggt ggt cgc      336
Asn His Gln Gly Phe Tyr Phe Ala Val Phe Pro Ser Val Gly Gly Arg
            100                 105                 110

cag ttc gaa gct gat aat atc gat cag atg gaa gcg gtt ggg cgt tat      384
Gln Phe Glu Ala Asp Asn Ile Asp Gln Met Glu Ala Val Gly Arg Tyr
        115                 120                 125

tta ggg cgt atg cac cag acg ggg cgc aaa cag ctt ttt atc cat cgc      432
Leu Gly Arg Met His Gln Thr Gly Arg Lys Gln Leu Phe Ile His Arg
    130                 135                 140

ccg acc atc ggt ttg aac gaa tat ctc att gag cca cgc aag ctg ttt      480
Pro Thr Ile Gly Leu Asn Glu Tyr Leu Ile Glu Pro Arg Lys Leu Phe
145                 150                 155                 160

gag gac gct aca ctg ata cct tcc ggg ttg aaa gcg gca ttc ctg aaa      528
Glu Asp Ala Thr Leu Ile Pro Ser Gly Leu Lys Ala Ala Phe Leu Lys
                165                 170                 175

gcg aca gat gag ctg att gcc gcc gtt aca gca cac tgg cgg gaa gat      576
Ala Thr Asp Glu Leu Ile Ala Ala Val Thr Ala His Trp Arg Glu Asp
            180                 185                 190

ttc acc gtt ctg cgg cta cat gga gac tgc cac gcc ggg aat att ctc      624
Phe Thr Val Leu Arg Leu His Gly Asp Cys His Ala Gly Asn Ile Leu
        195                 200                 205

tgg cgc gat ggt cca atg ttt gtt gat ctg gat gat gca cgt aat ggt      672
Trp Arg Asp Gly Pro Met Phe Val Asp Leu Asp Asp Ala Arg Asn Gly
    210                 215                 220

cca gcc gtt cag gat ttg tgg atg ttg ctc aat ggc gat aaa gcc gag      720
Pro Ala Val Gln Asp Leu Trp Met Leu Leu Asn Gly Asp Lys Ala Glu
225                 230                 235                 240

cag cgg atg caa ctg gaa act att att gaa gct tat gaa gaa ttt agc      768
Gln Arg Met Gln Leu Glu Thr Ile Ile Glu Ala Tyr Glu Glu Phe Ser
                245                 250                 255

gag ttc gac acc gct gaa atc gga ctg att gaa cct tta cgc gcc atg      816
Glu Phe Asp Thr Ala Glu Ile Gly Leu Ile Glu Pro Leu Arg Ala Met
            260                 265                 270

cgt ttg gtt tat tat ctt gcc tgg cta atg cgg cgt tgg gct gat ccc      864
Arg Leu Val Tyr Tyr Leu Ala Trp Leu Met Arg Arg Trp Ala Asp Pro
        275                 280                 285

gcg ttc ccg aaa aat ttc ccg tgg tta acc ggg gaa gat tac tgg ctg      912
Ala Phe Pro Lys Asn Phe Pro Trp Leu Thr Gly Glu Asp Tyr Trp Leu
    290                 295                 300

cga cag acg gcg act ttt ata gaa cag gca aaa gtt cta caa gaa ccc      960
```

```
Arg Gln Thr Ala Thr Phe Ile Glu Gln Ala Lys Val Leu Gln Glu Pro
305                 310                 315                 320

cct ttg caa tta aca cct atg tat taa                              987
Pro Leu Gln Leu Thr Pro Met Tyr  *
                325


<210> 226
<211> 627
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(627)

<400> 226
atg aaa aag att tgg ctg gcg ctg gct ggt tta gtt tta gcg ttt agc   48
Met Lys Lys Ile Trp Leu Ala Leu Ala Gly Leu Val Leu Ala Phe Ser
 1               5                  10                  15

gca tcg gcg gcg cag tat gaa gat ggt aaa cag tac act acc ctg gaa   96
Ala Ser Ala Ala Gln Tyr Glu Asp Gly Lys Gln Tyr Thr Thr Leu Glu
            20                  25                  30

aaa ccg gta gct ggc gcg ccg caa gtg ctg gag ttt ttc tct ttc ttc  144
Lys Pro Val Ala Gly Ala Pro Gln Val Leu Glu Phe Phe Ser Phe Phe
        35                  40                  45

tgc ccg cac tgc tat cag ttt gaa gaa gtt ctg cat att tct gat aat  192
Cys Pro His Cys Tyr Gln Phe Glu Glu Val Leu His Ile Ser Asp Asn
    50                  55                  60

gtg aag aaa aaa ctg ccg gaa ggc gtg aag atg act aaa tac cac gtc  240
Val Lys Lys Lys Leu Pro Glu Gly Val Lys Met Thr Lys Tyr His Val
65                  70                  75                  80

aac ttc atg ggt ggt gac ctg ggc aaa gat ctg act cag gca tgg gct  288
Asn Phe Met Gly Gly Asp Leu Gly Lys Asp Leu Thr Gln Ala Trp Ala
                85                  90                  95

gtg gcg atg gcg ctg ggc gtg gaa gac aaa gtg act gtt ccg ctg ttt  336
Val Ala Met Ala Leu Gly Val Glu Asp Lys Val Thr Val Pro Leu Phe
            100                 105                 110

gaa ggc gta cag aaa acc cag acc att cgt tct gct tct gat atc cgc  384
Glu Gly Val Gln Lys Thr Gln Thr Ile Arg Ser Ala Ser Asp Ile Arg
        115                 120                 125

gat gta ttt atc aac gca ggt att aaa ggt gaa gag tac gac gcg gcg  432
Asp Val Phe Ile Asn Ala Gly Ile Lys Gly Glu Glu Tyr Asp Ala Ala
    130                 135                 140

tgg aac agc ttc gtg gtg aaa tct ctg gtc gct cag cag gaa aaa gct  480
Trp Asn Ser Phe Val Val Lys Ser Leu Val Ala Gln Gln Glu Lys Ala
145                 150                 155                 160

gca gct gac gtg caa ttg cgt ggc gtt ccg gcg atg ttt gtt aac ggt  528
Ala Ala Asp Val Gln Leu Arg Gly Val Pro Ala Met Phe Val Asn Gly
                165                 170                 175
```

aaa tat cag ctg aat ccg cag ggt atg gat acc agc aat atg gat gtt 576
Lys Tyr Gln Leu Asn Pro Gln Gly Met Asp Thr Ser Asn Met Asp Val
180 185 190

ttt gtt cag cag tat gct gat aca gtg aaa tat ctg tcc gag aaa aaa 624
Phe Val Gln Gln Tyr Ala Asp Thr Val Lys Tyr Leu Ser Glu Lys Lys
195 200 205

taa 627
*

<210> 227
<211> 339
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(339)

<400> 227
atg acg atg aat tca ttc gag cga agg aat aag atc atc caa tta gtg 48
Met Thr Met Asn Ser Phe Glu Arg Arg Asn Lys Ile Ile Gln Leu Val
1 5 10 15

aat gaa cag gga acc gtg ctt gtt cag gat ctg gcg gga gta ttt gct 96
Asn Glu Gln Gly Thr Val Leu Val Gln Asp Leu Ala Gly Val Phe Ala
20 25 30

gcc tcg gaa gcg aca atc cgt gcc gat ttg cgc ttt ctc gaa caa aaa 144
Ala Ser Glu Ala Thr Ile Arg Ala Asp Leu Arg Phe Leu Glu Gln Lys
35 40 45

ggc gtg gtt acg cgc ttt cat ggc ggt gcg gcg aaa ata atg tct ggt 192
Gly Val Val Thr Arg Phe His Gly Gly Ala Ala Lys Ile Met Ser Gly
50 55 60

aat agt gaa acc gag acc cag gaa gtc ggg ttt aaa gag cga ttt cag 240
Asn Ser Glu Thr Glu Thr Gln Glu Val Gly Phe Lys Glu Arg Phe Gln
65 70 75 80

ctc gcc agc gcg cca aaa aac aga ata gcg cag gcg gca gtc aaa atg 288
Leu Ala Ser Ala Pro Lys Asn Arg Ile Ala Gln Ala Ala Val Lys Met
85 90 95

atc cac gaa ggg atg act gat cct acc cac gta ata tgg aca cag gcc 336
Ile His Glu Gly Met Thr Asp Pro Thr His Val Ile Trp Thr Gln Ala
100 105 110

taa 339
*

<210> 228
<211> 1041
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1041)

<400> 228
atg aaa tca gtg gtg aat gat act gat ggt atc gtg cgc gtt gca gaa 48
Met Lys Ser Val Val Asn Asp Thr Asp Gly Ile Val Arg Val Ala Glu
1 5 10 15

agc gtc att cct gaa att aaa cat cag gat gag gtg cgg gta aaa att 96
Ser Val Ile Pro Glu Ile Lys His Gln Asp Glu Val Arg Val Lys Ile
20 25 30

gcc agc tcg ggc tta tgt ggt tcc gat tta ccc agg ata ttt aaa aat 144
Ala Ser Ser Gly Leu Cys Gly Ser Asp Leu Pro Arg Ile Phe Lys Asn
35 40 45

ggt gca cat tat tat cca ata acg tta ggc cat gaa ttt agc ggc tat 192
Gly Ala His Tyr Tyr Pro Ile Thr Leu Gly His Glu Phe Ser Gly Tyr
50 55 60

att gat gcg gtg gga tcc ggt gtt gat gat tta cac cct ggc gat gcg 240
Ile Asp Ala Val Gly Ser Gly Val Asp Asp Leu His Pro Gly Asp Ala
65 70 75 80

gtt gcc tgt gtg ccg tta tta ccc tgt ttt act tgt cca gag tgt ctg 288
Val Ala Cys Val Pro Leu Leu Pro Cys Phe Thr Cys Pro Glu Cys Leu
85 90 95

aaa ggg ttt tat tcc cag tgc gca aaa tat gat ttt att ggc tcg cgg 336
Lys Gly Phe Tyr Ser Gln Cys Ala Lys Tyr Asp Phe Ile Gly Ser Arg
100 105 110

cgt gat ggt gga ttt gct gaa tat att gtc gtt aag cga aaa aat gtc 384
Arg Asp Gly Gly Phe Ala Glu Tyr Ile Val Val Lys Arg Lys Asn Val
115 120 125

ttt gct cta ccc acg gat atg cct att gag gat ggg gct ttt att gag 432
Phe Ala Leu Pro Thr Asp Met Pro Ile Glu Asp Gly Ala Phe Ile Glu
130 135 140

ccg att acc gtt ggt ctg cat gct ttt cat tta gcg caa ggt tgt gag 480
Pro Ile Thr Val Gly Leu His Ala Phe His Leu Ala Gln Gly Cys Glu
145 150 155 160

aat aaa aac gtt att att att ggt gcc gga acc att ggc ctg ctg gcc 528
Asn Lys Asn Val Ile Ile Ile Gly Ala Gly Thr Ile Gly Leu Leu Ala
165 170 175

att cag tgc gct gtc gcg ctg gga gca aag agt gtg acg gcg atc gac 576
Ile Gln Cys Ala Val Ala Leu Gly Ala Lys Ser Val Thr Ala Ile Asp
180 185 190

att agt tca gaa aaa ctg gca ctg gca aaa tct ttc ggt gcg atg caa 624
Ile Ser Ser Glu Lys Leu Ala Leu Ala Lys Ser Phe Gly Ala Met Gln
195 200 205

aca ttt aac agt agc gaa atg agc gcg ccg caa atg cag agc gtt tta 672
Thr Phe Asn Ser Ser Glu Met Ser Ala Pro Gln Met Gln Ser Val Leu
210 215 220

```
cgc gaa ctg cgc ttt aat cag ctt atc ctc gag acg gct ggc gta ccg    720
Arg Glu Leu Arg Phe Asn Gln Leu Ile Leu Glu Thr Ala Gly Val Pro
225                 230                 235                 240

caa act gtc gaa ctg gcg gta gag att gcc ggt cct cat gcc caa ctg    768
Gln Thr Val Glu Leu Ala Val Glu Ile Ala Gly Pro His Ala Gln Leu
                245                 250                 255

gcg ctg gtg ggc acg ttg cat cag gat ctg cat tta aca tcg gca acg    816
Ala Leu Val Gly Thr Leu His Gln Asp Leu His Leu Thr Ser Ala Thr
            260                 265                 270

ttt ggc aaa ata ttg cgt aaa gag ctg acg gtt atc ggc agt tgg atg    864
Phe Gly Lys Ile Leu Arg Lys Glu Leu Thr Val Ile Gly Ser Trp Met
        275                 280                 285

aac tac tcc agc cct tgg ccg ggg cag gag tgg gaa acg gcg agc cgg    912
Asn Tyr Ser Ser Pro Trp Pro Gly Gln Glu Trp Glu Thr Ala Ser Arg
    290                 295                 300

ttg ctg aca gaa cgt aag tta agc ctg gag cca tta atc gct cac cgt    960
Leu Leu Thr Glu Arg Lys Leu Ser Leu Glu Pro Leu Ile Ala His Arg
305                 310                 315                 320

gga agc ttt gaa agc ttc gcc cag gcg gtg cgt gac atc gct cgt aat    1008
Gly Ser Phe Glu Ser Phe Ala Gln Ala Val Arg Asp Ile Ala Arg Asn
                325                 330                 335

gct atg ccg ggc aaa gtg ttg ctc att ccc tga                         1041
Ala Met Pro Gly Lys Val Leu Leu Ile Pro  *
            340                 345


<210> 229
<211> 1356
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1356)

<400> 229
atg ttt tca gaa gtc atg cgt tat att ctc gac ctc ggc cct acg gtg    48
Met Phe Ser Glu Val Met Arg Tyr Ile Leu Asp Leu Gly Pro Thr Val
 1               5                  10                  15

atg ctg ccg att gtc atc att att ttt tct aaa ata tta ggc atg aag    96
Met Leu Pro Ile Val Ile Ile Ile Phe Ser Lys Ile Leu Gly Met Lys
            20                  25                  30

gca ggc gat tgc ttt aaa gcg ggt ctg cat atc ggg att ggc ttt gtt    144
Ala Gly Asp Cys Phe Lys Ala Gly Leu His Ile Gly Ile Gly Phe Val
        35                  40                  45

ggc att ggc ctt gtg att ggc tta atg ctg gat tcc att ggc ccg gcg    192
Gly Ile Gly Leu Val Ile Gly Leu Met Leu Asp Ser Ile Gly Pro Ala
    50                  55                  60

gcg aaa gcg atg gcg gaa aat ttc gac ctg aat ctg cat gtg gtc gat    240
```

```
Ala Lys Ala Met Ala Glu Asn Phe Asp Leu Asn Leu His Val Val Asp
 65                 70                  75                  80

gtt ggc tgg ccg ggc tct tca cca atg acc tgg gcg tcg caa att gcg    288
Val Gly Trp Pro Gly Ser Ser Pro Met Thr Trp Ala Ser Gln Ile Ala
                85                  90                  95

ctg gtg gcg att ccg att gcg att ctg gtt aac gtg gcg atg tta ctg    336
Leu Val Ala Ile Pro Ile Ala Ile Leu Val Asn Val Ala Met Leu Leu
            100                 105                 110

acc cgt atg acg cgg gtg gta aat gtt gat atc tgg aat atc tgg cat    384
Thr Arg Met Thr Arg Val Val Asn Val Asp Ile Trp Asn Ile Trp His
        115                 120                 125

atg acc ttc acc ggc gcg ttg ctg cat ctg gca acc ggt tca tgg atg    432
Met Thr Phe Thr Gly Ala Leu Leu His Leu Ala Thr Gly Ser Trp Met
    130                 135                 140

ata ggg atg gca ggt gtg gta att cac gcg gcg ttt gtt tat aag ctc    480
Ile Gly Met Ala Gly Val Val Ile His Ala Ala Phe Val Tyr Lys Leu
145                 150                 155                 160

ggc gac tgg ttt gcc cgc gat acc cga aat ttc ttt gag ctg gaa ggt    528
Gly Asp Trp Phe Ala Arg Asp Thr Arg Asn Phe Phe Glu Leu Glu Gly
                165                 170                 175

att gct att ccg cac ggt acg tcg gcg tat atg ggg ccg att gcg gtg    576
Ile Ala Ile Pro His Gly Thr Ser Ala Tyr Met Gly Pro Ile Ala Val
            180                 185                 190

ctg gtc gat gct atc atc gag aaa atc cca ggc gtt aac cga att aaa    624
Leu Val Asp Ala Ile Ile Glu Lys Ile Pro Gly Val Asn Arg Ile Lys
        195                 200                 205

ttt agc gcc gac gat att cag cgc aaa ttt ggt cca ttt ggc gag cct    672
Phe Ser Ala Asp Asp Ile Gln Arg Lys Phe Gly Pro Phe Gly Glu Pro
    210                 215                 220

gtc acc gtg ggt ttt gtg atg ggg ctg att atc ggc atc ctc gcg ggt    720
Val Thr Val Gly Phe Val Met Gly Leu Ile Ile Gly Ile Leu Ala Gly
225                 230                 235                 240

tac gat gtc aaa ggt gta ttg cag ctg gcg gta aaa acg gcg gca gtg    768
Tyr Asp Val Lys Gly Val Leu Gln Leu Ala Val Lys Thr Ala Ala Val
                245                 250                 255

atg ctg cta atg cca cgg gtg att aaa ccc atc atg gat ggt tta acg    816
Met Leu Leu Met Pro Arg Val Ile Lys Pro Ile Met Asp Gly Leu Thr
            260                 265                 270

ccc atc gct aag cag gct cgt agt cgt tta cag gcg aag ttc ggc ggt    864
Pro Ile Ala Lys Gln Ala Arg Ser Arg Leu Gln Ala Lys Phe Gly Gly
        275                 280                 285

cag gag ttc ctg att ggc ctt gat ccg gcg ttg ctg ctg gga cat acg    912
Gln Glu Phe Leu Ile Gly Leu Asp Pro Ala Leu Leu Leu Gly His Thr
    290                 295                 300

gcg gtg gta tcg gca agc ctg att ttt atc cca ctc acc att tta att    960
Ala Val Val Ser Ala Ser Leu Ile Phe Ile Pro Leu Thr Ile Leu Ile
```

```
305                 310                 315                 320

gct gtt tgt gtg ccg ggt aat cag gtg ctg ccg ttt ggc gat ctt gcc     1008
Ala Val Cys Val Pro Gly Asn Gln Val Leu Pro Phe Gly Asp Leu Ala
                325                 330                 335

acc atc ggc ttc ttc gtg gcg atg gcg gtc gcc gtg cat cgt gga aat     1056
Thr Ile Gly Phe Phe Val Ala Met Ala Val Ala Val His Arg Gly Asn
            340                 345                 350

ctg ttc cgc acc tta atc tcg ggt gtc atc att atg agc atc acc ctg     1104
Leu Phe Arg Thr Leu Ile Ser Gly Val Ile Ile Met Ser Ile Thr Leu
        355                 360                 365

tgg atc gcg acg caa act att ggt ttg cac acc caa ctg gcg gct aat     1152
Trp Ile Ala Thr Gln Thr Ile Gly Leu His Thr Gln Leu Ala Ala Asn
    370                 375                 380

gct ggg gcg tta aaa gcc ggg ggt atg gtg gct tca atg gat cag ggc     1200
Ala Gly Ala Leu Lys Ala Gly Gly Met Val Ala Ser Met Asp Gln Gly
385                 390                 395                 400

ggt tct ccc att acc tgg tta ctg att cag gtt ttc tcc ccg caa aat     1248
Gly Ser Pro Ile Thr Trp Leu Leu Ile Gln Val Phe Ser Pro Gln Asn
                405                 410                 415

att ccc ggt ttc att att atc ggt gca att tat ctg acc ggt att ttc     1296
Ile Pro Gly Phe Ile Ile Ile Gly Ala Ile Tyr Leu Thr Gly Ile Phe
            420                 425                 430

atg acc tgg cgt aga gcg cgt ggc ttt att aaa caa gag aaa gtc gtt     1344
Met Thr Trp Arg Arg Ala Arg Gly Phe Ile Lys Gln Glu Lys Val Val
        435                 440                 445

ctc gca gaa taa                                                     1356
Leu Ala Glu  *
    450


<210> 230
<211> 285
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(285)

<400> 230
atg aaa cgc aag att att gtc gct tgc gga ggc gcg gtt gcg acc tct     48
Met Lys Arg Lys Ile Ile Val Ala Cys Gly Gly Ala Val Ala Thr Ser
 1               5                  10                  15

acg atg gcg gcg gaa gaa att aaa gag ttg tgt cag aat cat aat att     96
Thr Met Ala Ala Glu Glu Ile Lys Glu Leu Cys Gln Asn His Asn Ile
            20                  25                  30

cct gtt gaa tta atc cag tgt cgg gtt aat gaa ata gaa acc tat atg     144
Pro Val Glu Leu Ile Gln Cys Arg Val Asn Glu Ile Glu Thr Tyr Met
        35                  40                  45
```

```
gat ggt gtg cat ttg ata tgc acc act gcc aaa gtg gat cgt agt ttt    192
Asp Gly Val His Leu Ile Cys Thr Thr Ala Lys Val Asp Arg Ser Phe
     50                  55                  60

ggc gat att ccg tta gtt cac ggc atg cct ttt att tct ggt atc ggt    240
Gly Asp Ile Pro Leu Val His Gly Met Pro Phe Ile Ser Gly Ile Gly
 65                  70                  75                  80

atc gaa gca tta caa aat aaa att ctg act atc tta cag ggg tga        285
Ile Glu Ala Leu Gln Asn Lys Ile Leu Thr Ile Leu Gln Gly  *
                 85                  90
```

<210> 231
<211> 453
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(453)

<400> 231

```
atg act aac ctg ttt gtt cgt agc gga att tct ttt gtc gat cgt agc    48
Met Thr Asn Leu Phe Val Arg Ser Gly Ile Ser Phe Val Asp Arg Ser
 1               5                  10                  15

gaa gtt tta acc cat atc ggt aat gag atg ctc gcc aaa ggt gtg gtt    96
Glu Val Leu Thr His Ile Gly Asn Glu Met Leu Ala Lys Gly Val Val
             20                  25                  30

cat gat acc tgg cca cag gca tta att gcc aga gaa gca gaa ttc cct    144
His Asp Thr Trp Pro Gln Ala Leu Ile Ala Arg Glu Ala Glu Phe Pro
         35                  40                  45

acc ggg ata atg ctt gag cag cac gct att gca ata ccg cat tgt gag    192
Thr Gly Ile Met Leu Glu Gln His Ala Ile Ala Ile Pro His Cys Glu
     50                  55                  60

gcg att cat gct aag tcg tca gcc att tat ctg tta agg cca aca aat    240
Ala Ile His Ala Lys Ser Ser Ala Ile Tyr Leu Leu Arg Pro Thr Asn
 65                  70                  75                  80

aaa gtt cat ttt cag caa gcg gat gat gat aac gac gtg gcg gta tcg    288
Lys Val His Phe Gln Gln Ala Asp Asp Asp Asn Asp Val Ala Val Ser
                 85                  90                  95

ttg gtt att gcg ttg att gtg gaa aat ccg cag cag caa ttg aaa ctt    336
Leu Val Ile Ala Leu Ile Val Glu Asn Pro Gln Gln Gln Leu Lys Leu
            100                 105                 110

tta cgc tgt tta ttt ggc aag tta caa cag ccc gat atc gtc gag aca    384
Leu Arg Cys Leu Phe Gly Lys Leu Gln Gln Pro Asp Ile Val Glu Thr
        115                 120                 125

cta atc act ctt cct gaa acc cag tta aag gaa tac ttc aca aag tat    432
Leu Ile Thr Leu Pro Glu Thr Gln Leu Lys Glu Tyr Phe Thr Lys Tyr
    130                 135                 140

gtt tta gat tca gac gaa taa                                        453
```

Val Leu Asp Ser Asp Glu *
145 150

<210> 232
<211> 1263
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1263)

<400> 232
atg aaa acg tta att gcc cgg cat aaa gct ggt gaa cat atc ggc ata 48
Met Lys Thr Leu Ile Ala Arg His Lys Ala Gly Glu His Ile Gly Ile
1 5 10 15

tgt tca gtc tgt tct gcc cat ccg ttg gtt atc gaa gcg gcg ctg gca 96
Cys Ser Val Cys Ser Ala His Pro Leu Val Ile Glu Ala Ala Leu Ala
20 25 30

ttt gat cgc aac agc acg cgc aaa gtg ctg att gaa gca acg tca aac 144
Phe Asp Arg Asn Ser Thr Arg Lys Val Leu Ile Glu Ala Thr Ser Asn
35 40 45

cag gtc aat caa ttt ggc ggt tat acc gga atg aca ccg gca gac ttt 192
Gln Val Asn Gln Phe Gly Gly Tyr Thr Gly Met Thr Pro Ala Asp Phe
50 55 60

cgc gaa ttt gtt ttt acg att gcc gat aaa gtt ggg ttt gca cgc gaa 240
Arg Glu Phe Val Phe Thr Ile Ala Asp Lys Val Gly Phe Ala Arg Glu
65 70 75 80

cgc att att ctc ggc ggc gat cat ctg ggg cca aac tgc tgg cag caa 288
Arg Ile Ile Leu Gly Gly Asp His Leu Gly Pro Asn Cys Trp Gln Gln
85 90 95

gaa aat gcg gat gcg gcg atg gaa aaa tcc gtc gag ctg gta aag gaa 336
Glu Asn Ala Asp Ala Ala Met Glu Lys Ser Val Glu Leu Val Lys Glu
100 105 110

tat gtt cgt gcc ggc ttc agt aaa att cat ctt gat gcg tca atg tcc 384
Tyr Val Arg Ala Gly Phe Ser Lys Ile His Leu Asp Ala Ser Met Ser
115 120 125

tgc gcg ggg gat ccc ata ccg tta gca cca gaa acg gtt gcg gaa cga 432
Cys Ala Gly Asp Pro Ile Pro Leu Ala Pro Glu Thr Val Ala Glu Arg
130 135 140

gct gct gtg ctt tgc ttt gct gcg gaa agt gtg gcg aca gat tgc cag 480
Ala Ala Val Leu Cys Phe Ala Ala Glu Ser Val Ala Thr Asp Cys Gln
145 150 155 160

cgt gag caa ctg agc tat gtc att ggc acc gaa gtt ccg gtt ccg ggc 528
Arg Glu Gln Leu Ser Tyr Val Ile Gly Thr Glu Val Pro Val Pro Gly
165 170 175

ggt gag gcc agc gcc att cag tca gta cac atc acc cat gtt gaa gat 576
Gly Glu Ala Ser Ala Ile Gln Ser Val His Ile Thr His Val Glu Asp
180 185 190

gcc gcc aat act tta cgt acg cat caa aag gcc ttt att gcc cgt ggg 624
Ala Ala Asn Thr Leu Arg Thr His Gln Lys Ala Phe Ile Ala Arg Gly
195 200 205

ctg aca gag gcg tta aca cgt gtg att gcc atc gtg gtg cag ccg ggt 672
Leu Thr Glu Ala Leu Thr Arg Val Ile Ala Ile Val Val Gln Pro Gly
210 215 220

gtg gaa ttt gat cac agc aat att atc cat tat cag ccg cag gaa gcg 720
Val Glu Phe Asp His Ser Asn Ile Ile His Tyr Gln Pro Gln Glu Ala
225 230 235 240

cag ccg ctg gcg caa tgg ata gaa aac acc cga atg gtt tat gaa gca 768
Gln Pro Leu Ala Gln Trp Ile Glu Asn Thr Arg Met Val Tyr Glu Ala
245 250 255

cat tct acc gat tac cag acc cgg acg gct tat tgg gaa tta gtc cgc 816
His Ser Thr Asp Tyr Gln Thr Arg Thr Ala Tyr Trp Glu Leu Val Arg
260 265 270

gat cac ttt gca ata ttg aaa gtc ggt ccc gca tta acc ttt gct tta 864
Asp His Phe Ala Ile Leu Lys Val Gly Pro Ala Leu Thr Phe Ala Leu
275 280 285

cgc gag gcg ata ttt gca ctg gca caa att gag cag gaa ctt atc gcc 912
Arg Glu Ala Ile Phe Ala Leu Ala Gln Ile Glu Gln Glu Leu Ile Ala
290 295 300

cct gaa aat cgc agc ggt tgc ctg gcg gta att gaa gaa gtg atg ctc 960
Pro Glu Asn Arg Ser Gly Cys Leu Ala Val Ile Glu Glu Val Met Leu
305 310 315 320

gac gaa ccg caa tac tgg aaa aaa tat tat cgt acg ggt ttt aac gat 1008
Asp Glu Pro Gln Tyr Trp Lys Lys Tyr Tyr Arg Thr Gly Phe Asn Asp
325 330 335

tca tta ctg gat att cgt tac agc ctg tcg gat cgt att cgt tat tac 1056
Ser Leu Leu Asp Ile Arg Tyr Ser Leu Ser Asp Arg Ile Arg Tyr Tyr
340 345 350

tgg ccg cat agt cgg att aaa aat agc gtc gaa acg atg atg gtg aat 1104
Trp Pro His Ser Arg Ile Lys Asn Ser Val Glu Thr Met Met Val Asn
355 360 365

ctt gaa ggc gtg gac atc cca ctg ggc atg att agt cag tat ctt ccc 1152
Leu Glu Gly Val Asp Ile Pro Leu Gly Met Ile Ser Gln Tyr Leu Pro
370 375 380

aaa caa ttt gaa cgc att cag tcc ggg gaa tta tca gca ata ccg cat 1200
Lys Gln Phe Glu Arg Ile Gln Ser Gly Glu Leu Ser Ala Ile Pro His
385 390 395 400

cag ctg att atg gat aaa att tat gat gtt ttg cgc gcc tat cgc tac 1248
Gln Leu Ile Met Asp Lys Ile Tyr Asp Val Leu Arg Ala Tyr Arg Tyr
405 410 415

ggc tgt gcg gaa taa 1263
Gly Cys Ala Glu *
420

<210> 233
<211> 861
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(861)

<400> 233
atg aaa atg tac gtg gta tcg aca aag cag atg ctg aac aac gca cag 48
Met Lys Met Tyr Val Val Ser Thr Lys Gln Met Leu Asn Asn Ala Gln
1 5 10 15

cgc ggc ggt tat gcg gtt ccg gca ttc aat att cac aat ctc gaa acg 96
Arg Gly Gly Tyr Ala Val Pro Ala Phe Asn Ile His Asn Leu Glu Thr
20 25 30

atg caa gtg gtg gta gaa acc gct gcc aac ctg cat gcg ccg gtc atc 144
Met Gln Val Val Val Glu Thr Ala Ala Asn Leu His Ala Pro Val Ile
35 40 45

atc gcc gga acg cct ggc aca ttt act cat gct ggt aca gaa aat ctg 192
Ile Ala Gly Thr Pro Gly Thr Phe Thr His Ala Gly Thr Glu Asn Leu
50 55 60

ttg gcg ctg gtc agc gcg atg gcg aag caa tat cac cat cca ctg gca 240
Leu Ala Leu Val Ser Ala Met Ala Lys Gln Tyr His His Pro Leu Ala
65 70 75 80

att cat ctc gac cat cac acg aaa ttt gac gat atc gct cag aag gtt 288
Ile His Leu Asp His His Thr Lys Phe Asp Asp Ile Ala Gln Lys Val
85 90 95

cgt tct ggc gtg cgc tca gtc atg att gac gcc tcg cat ttg cct ttt 336
Arg Ser Gly Val Arg Ser Val Met Ile Asp Ala Ser His Leu Pro Phe
100 105 110

gcg caa aat att tca cgg gtc aaa gag gtg gtg gat ttt tgc cat cgc 384
Ala Gln Asn Ile Ser Arg Val Lys Glu Val Val Asp Phe Cys His Arg
115 120 125

ttt gat gtc agc gtc gag gcg gag ctg ggg caa ctt ggc ggc cag gaa 432
Phe Asp Val Ser Val Glu Ala Glu Leu Gly Gln Leu Gly Gly Gln Glu
130 135 140

gat gat gtg caa gtc aat gaa gcc gat gcg ttg tac acc aac ccc gct 480
Asp Asp Val Gln Val Asn Glu Ala Asp Ala Leu Tyr Thr Asn Pro Ala
145 150 155 160

cag gcg cgt gaa ttt gcc gag gca acc gga att gat tcc ctg gcg gtc 528
Gln Ala Arg Glu Phe Ala Glu Ala Thr Gly Ile Asp Ser Leu Ala Val
165 170 175

gcc atc ggc acg gct cat ggg atg tat gcc agc gca ccg gcg ctt gat 576
Ala Ile Gly Thr Ala His Gly Met Tyr Ala Ser Ala Pro Ala Leu Asp
180 185 190

ttt tct aga ctg gag aac att cgc cag tgg gtg aac tta ccg ctg gtg 624
Phe Ser Arg Leu Glu Asn Ile Arg Gln Trp Val Asn Leu Pro Leu Val

```
        195                 200                 205

ctg cat ggc gcg tca ggg tta tcg act aag gat att cag caa acc atc    672
Leu His Gly Ala Ser Gly Leu Ser Thr Lys Asp Ile Gln Gln Thr Ile
    210                 215                 220

aaa ctg ggg ata tgc aaa atc aac gtt gca acg gag ctg aaa aat gcc    720
Lys Leu Gly Ile Cys Lys Ile Asn Val Ala Thr Glu Leu Lys Asn Ala
225                 230                 235                 240

ttc tcg cag gcg tta aaa aat tac ctg acc gag cac cct gaa gcg acc    768
Phe Ser Gln Ala Leu Lys Asn Tyr Leu Thr Glu His Pro Glu Ala Thr
                245                 250                 255

gat ccc cgg gat tat ttg cag tcg gct aaa tcc gca atg cgc gat gtg    816
Asp Pro Arg Asp Tyr Leu Gln Ser Ala Lys Ser Ala Met Arg Asp Val
            260                 265                 270

gtg agc aaa gtg att gcc gat tgt ggc tgc gag ggc agg gca taa        861
Val Ser Lys Val Ile Ala Asp Cys Gly Cys Glu Gly Arg Ala  *
        275                 280                 285



<210> 234
<211> 474
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(474)

<400> 234
atg agt caa aac gat atc att atc aga act cat tat aag tct cct cat    48
Met Ser Gln Asn Asp Ile Ile Ile Arg Thr His Tyr Lys Ser Pro His
 1               5                  10                  15

aga ttg cac atc gat agc gac ata cca aca cct tca tca gag cct att    96
Arg Leu His Ile Asp Ser Asp Ile Pro Thr Pro Ser Ser Glu Pro Ile
            20                  25                  30

aat caa ttt gcg cgc cag ctc atc acc cta ctt gat acc tct gac tta    144
Asn Gln Phe Ala Arg Gln Leu Ile Thr Leu Leu Asp Thr Ser Asp Leu
        35                  40                  45

agt tcg atg ctg tca tac tgt gtt act cag gaa ttt acc gca aac tgt    192
Ser Ser Met Leu Ser Tyr Cys Val Thr Gln Glu Phe Thr Ala Asn Cys
    50                  55                  60

cga aaa ata tca caa aat tgt tat tcc act gcc ctt ttt acc att aac    240
Arg Lys Ile Ser Gln Asn Cys Tyr Ser Thr Ala Leu Phe Thr Ile Asn
65                  70                  75                  80

ttt gcc act tca ccc atc cat aca gaa aat ata ctc att aca tta cac    288
Phe Ala Thr Ser Pro Ile His Thr Glu Asn Ile Leu Ile Thr Leu His
                85                  90                  95

tat aaa aaa gaa atc att tcc tta tta ctg gaa acc acg cct att aaa    336
Tyr Lys Lys Glu Ile Ile Ser Leu Leu Leu Glu Thr Thr Pro Ile Lys
            100                 105                 110
```

gct aac cat ttg cga agc ata ctg gat tat att gaa cag gaa cag tta 384
Ala Asn His Leu Arg Ser Ile Leu Asp Tyr Ile Glu Gln Glu Gln Leu
115 120 125

act gct gaa gat cgt aac cat tgt atg aaa ctg tct aaa aaa atc cat 432
Thr Ala Glu Asp Arg Asn His Cys Met Lys Leu Ser Lys Lys Ile His
130 135 140

aga gaa aaa aac tat aca ccc aac agt aaa tct caa tgg tag 474
Arg Glu Lys Asn Tyr Thr Pro Asn Ser Lys Ser Gln Trp *
145 150 155

<210> 235
<211> 1416
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1416)

<400> 235
ttg caa tct cct tct gac gct att ttt tgt cgc cat ctg tca ttg caa 48
Met Gln Ser Pro Ser Asp Ala Ile Phe Cys Arg His Leu Ser Leu Gln
1 5 10 15

tac gcc ctt gat tca ttg aga aat gga aaa ggc aaa gtc aac ctg att 96
Tyr Ala Leu Asp Ser Leu Arg Asn Gly Lys Gly Lys Val Asn Leu Ile
20 25 30

aaa cat tac tcc tcc gtt gaa tcc ata cag cag cat gtc ccc tta gtc 144
Lys His Tyr Ser Ser Val Glu Ser Ile Gln Gln His Val Pro Leu Val
35 40 45

cgg gac gcg gag ttc aga gca tta ctt cgc cat cct cct gca ggg agt 192
Arg Asp Ala Glu Phe Arg Ala Leu Leu Arg His Pro Pro Ala Gly Ser
50 55 60

cgc gtt atc gcg agt aag gat ttt ggc ttc gct tta gat att ttc ttt 240
Arg Val Ile Ala Ser Lys Asp Phe Gly Phe Ala Leu Asp Ile Phe Phe
65 70 75 80

tgt cga atg atg gca aac aat gtc agt cat atg tcc gcg att tta tat 288
Cys Arg Met Met Ala Asn Asn Val Ser His Met Ser Ala Ile Leu Tyr
85 90 95

ata gac aat cat act ttg tca gta agg cta cga ata aag cag tca gtg 336
Ile Asp Asn His Thr Leu Ser Val Arg Leu Arg Ile Lys Gln Ser Val
100 105 110

tat gga caa tta aat tat gtt gtg tcc gtt tac gac ccg aac gat acc 384
Tyr Gly Gln Leu Asn Tyr Val Val Ser Val Tyr Asp Pro Asn Asp Thr
115 120 125

aac gtt gcc gtc aga gac acc cac agg aca gca cgg ggc ttt ctt tcg 432
Asn Val Ala Val Arg Asp Thr His Arg Thr Ala Arg Gly Phe Leu Ser
130 135 140

ctt gat aag ttc atc agt tca ggt ccc gat gct cag acc tgg gct gat 480
Leu Asp Lys Phe Ile Ser Ser Gly Pro Asp Ala Gln Thr Trp Ala Asp
145 150 155 160

aga tat gtt cgc aac tgt gca att gct att ctg ccc cta tta cct gtg 528
Arg Tyr Val Arg Asn Cys Ala Ile Ala Ile Leu Pro Leu Leu Pro Val
165 170 175

gga gtt cca ggg gct att ttc gcg ggt att gca tca cga atg cca ttt 576
Gly Val Pro Gly Ala Ile Phe Ala Gly Ile Ala Ser Arg Met Pro Phe
180 185 190

gcc cct ata cat cca tcg gca atg ttg tta ata atg gct aca ggc cag 624
Ala Pro Ile His Pro Ser Ala Met Leu Leu Ile Met Ala Thr Gly Gln
195 200 205

tct caa cag ctt att aca tta ttc aaa cag ttg ccc ata ctc cct gaa 672
Ser Gln Gln Leu Ile Thr Leu Phe Lys Gln Leu Pro Ile Leu Pro Glu
210 215 220

aaa gaa atc att gaa ata ata act gcg cag aat agc gtt ggt aca cct 720
Lys Glu Ile Ile Glu Ile Ile Thr Ala Gln Asn Ser Val Gly Thr Pro
225 230 235 240

gct tta ttt ttg gct atg atg aac gga cat act gac aac gtg aaa ata 768
Ala Leu Phe Leu Ala Met Met Asn Gly His Thr Asp Asn Val Lys Ile
245 250 255

ttt atg caa gaa att cag tca ctg gta gat aat cac atc att cat gaa 816
Phe Met Gln Glu Ile Gln Ser Leu Val Asp Asn His Ile Ile His Glu
260 265 270

gat aat ctg gtt aaa tta ctg caa act aaa agt gct aac gaa aca cct 864
Asp Asn Leu Val Lys Leu Leu Gln Thr Lys Ser Ala Asn Glu Thr Pro
275 280 285

gga ctt tat atc tcc atg ttg tat gga ttc gat gaa ata atc gat atc 912
Gly Leu Tyr Ile Ser Met Leu Tyr Gly Phe Asp Glu Ile Ile Asp Ile
290 295 300

ttt ctg aat gca tta acc act cca ata gca caa gag ctt tta aac aaa 960
Phe Leu Asn Ala Leu Thr Thr Pro Ile Ala Gln Glu Leu Leu Asn Lys
305 310 315 320

aaa ctg gtg atg agt att tta gcc atg aaa ata cat gat ggt gag cca 1008
Lys Leu Val Met Ser Ile Leu Ala Met Lys Ile His Asp Gly Glu Pro
325 330 335

gga tta tac gcc gca atg gaa aat aat cac cct ttg tgt gtc aca cgg 1056
Gly Leu Tyr Ala Ala Met Glu Asn Asn His Pro Leu Cys Val Thr Arg
340 345 350

ttc ctc tct aaa att aat ggc atc gcc ttt aaa tac aag ttg agc aaa 1104
Phe Leu Ser Lys Ile Asn Gly Ile Ala Phe Lys Tyr Lys Leu Ser Lys
355 360 365

gct aac atc atg gat tta tta aaa gga gct aca gca cag gga acc cct 1152
Ala Asn Ile Met Asp Leu Leu Lys Gly Ala Thr Ala Gln Gly Thr Pro
370 375 380

gca tta tac atc gcc atg agc aag ggt aat gaa gac gtc gtg tta tct 1200

```
Ala Leu Tyr Ile Ala Met Ser Lys Gly Asn Glu Asp Val Val Leu Ser
385                 390                 395                 400

tat ata tcg acg ctg ggt gct ttt gca aaa aaa cat tct ttt agt caa     1248
Tyr Ile Ser Thr Leu Gly Ala Phe Ala Lys Lys His Ser Phe Ser Gln
                405                 410                 415

cat cag tta ttt aca cta ttg gcc gct aaa aat cat gac aac atg tca     1296
His Gln Leu Phe Thr Leu Leu Ala Ala Lys Asn His Asp Asn Met Ser
            420                 425                 430

gct gtt cat ata gcc att cat cat aag cat tat aaa act gta gaa aca     1344
Ala Val His Ile Ala Ile His His Lys His Tyr Lys Thr Val Glu Thr
        435                 440                 445

tat tat gct gct att aat gca atc agc caa agc ctg agt ttt agt gct     1392
Tyr Tyr Ala Ala Ile Asn Ala Ile Ser Gln Ser Leu Ser Phe Ser Ala
    450                 455                 460

gat gaa ata aag acg tat tta taa                                     1416
Asp Glu Ile Lys Thr Tyr Leu  *
465                 470


<210> 236
<211> 387
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(387)

<400> 236
atg cca tct gga tta ttt atg gac tta ttg cct ttt tta ctg gac gcg     48
Met Pro Ser Gly Leu Phe Met Asp Leu Leu Pro Phe Leu Leu Asp Ala
 1               5                  10                  15

aac ctc agc gcg aca aac cca ccc gct att ccg cac tgg tgg aag cgt     96
Asn Leu Ser Ala Thr Asn Pro Pro Ala Ile Pro His Trp Trp Lys Arg
            20                  25                  30

caa ccg ctt att ccc aac ctt ctg tca cag gaa ctg aaa aac tat ctg     144
Gln Pro Leu Ile Pro Asn Leu Leu Ser Gln Glu Leu Lys Asn Tyr Leu
        35                  40                  45

aag ctt aat gtt aaa gag aaa aat att cag att gca gac cag gta att     192
Lys Leu Asn Val Lys Glu Lys Asn Ile Gln Ile Ala Asp Gln Val Ile
    50                  55                  60

att gat gaa act gca ggt gaa gtt gtt atc ggc gcg aat acc cgt att     240
Ile Asp Glu Thr Ala Gly Glu Val Val Ile Gly Ala Asn Thr Arg Ile
65                  70                  75                  80

tgt cat ggt gcc gtt att cag ggt ccg gta gtg att ggc gca aac tgc     288
Cys His Gly Ala Val Ile Gln Gly Pro Val Val Ile Gly Ala Asn Cys
                85                  90                  95

ctg ata ggt aat tgg att tgc ccc tat att tcc aga cat ctg tta tca     336
Leu Ile Gly Asn Trp Ile Cys Pro Tyr Ile Ser Arg His Leu Leu Ser
            100                 105                 110
```

```
ctt aac cca tta caa gcc cgc tgc cgc aga tat tcc cgt ggc gag cga     384
Leu Asn Pro Leu Gln Ala Arg Cys Arg Arg Tyr Ser Arg Gly Glu Arg
        115                 120                 125

taa                                                                 387
 *
```

<210> 237
<211> 1197
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1197)

<400> 237

```
atg aaa acc tgg ata ttt atc tgt atg tcc ata gca atg ttg cta tgg     48
Met Lys Thr Trp Ile Phe Ile Cys Met Ser Ile Ala Met Leu Leu Trp
 1               5                  10                  15

ttt tta agt acg cta aga cgt aaa ccc agt caa aag aaa ggc tgt att     96
Phe Leu Ser Thr Leu Arg Arg Lys Pro Ser Gln Lys Lys Gly Cys Ile
            20                  25                  30

gac gcc att ata cct gcg tat aac gaa ggc ccg tgt ctg gcg cag tca     144
Asp Ala Ile Ile Pro Ala Tyr Asn Glu Gly Pro Cys Leu Ala Gln Ser
        35                  40                  45

ctg gat aat cta ctg cgt aac cct tat ttt tgc cgg gta att tgc gtt     192
Leu Asp Asn Leu Leu Arg Asn Pro Tyr Phe Cys Arg Val Ile Cys Val
    50                  55                  60

aac gac ggc tcc acg gac aat acc gaa gcg gtc atg gcg gaa gtc aaa     240
Asn Asp Gly Ser Thr Asp Asn Thr Glu Ala Val Met Ala Glu Val Lys
65                  70                  75                  80

cgc aaa tgg ggc gac cgc ttt gtt gcc gtc acg caa aaa aat acc ggt     288
Arg Lys Trp Gly Asp Arg Phe Val Ala Val Thr Gln Lys Asn Thr Gly
                85                  90                  95

aaa ggt ggt gcg ctg atg aat ggc ctc aat tac gcc acc tgc gac cag     336
Lys Gly Gly Ala Leu Met Asn Gly Leu Asn Tyr Ala Thr Cys Asp Gln
            100                 105                 110

gtt ttt tta agt gat gcc gac acc tat gtt ccg ccc gat caa gac gga     384
Val Phe Leu Ser Asp Ala Asp Thr Tyr Val Pro Pro Asp Gln Asp Gly
        115                 120                 125

atg ggc tat atg ctg gca gaa att gag cgc ggt gcc gat gcc gta ggc     432
Met Gly Tyr Met Leu Ala Glu Ile Glu Arg Gly Ala Asp Ala Val Gly
    130                 135                 140

ggc att ccc tct act gcg ttg aaa ggc gcg ggt ctg tta ccg cac atc     480
Gly Ile Pro Ser Thr Ala Leu Lys Gly Ala Gly Leu Leu Pro His Ile
145                 150                 155                 160

cgc gcg acc gta aag ttg ccg atg att gtt atg aag cgc acg cta cag     528
```

```
Arg Ala Thr Val Lys Leu Pro Met Ile Val Met Lys Arg Thr Leu Gln
                165                 170                 175

cag ctc ctg ggt ggc gca ccg ttt att atc agc ggt gcc tgc ggg atg     576
Gln Leu Leu Gly Gly Ala Pro Phe Ile Ile Ser Gly Ala Cys Gly Met
            180                 185                 190

ttc cgt act gat gta ttg cgt aag ttc ggt ttc tcg gat cgt act aaa     624
Phe Arg Thr Asp Val Leu Arg Lys Phe Gly Phe Ser Asp Arg Thr Lys
        195                 200                 205

gtc gaa gac ctt gat ctc acc tgg aca ttg gtg gca aac ggc tac cgt     672
Val Glu Asp Leu Asp Leu Thr Trp Thr Leu Val Ala Asn Gly Tyr Arg
    210                 215                 220

att cgg cag gcg aat cgc tgc atc gta tac cca cag gaa tgc aac agc     720
Ile Arg Gln Ala Asn Arg Cys Ile Val Tyr Pro Gln Glu Cys Asn Ser
225                 230                 235                 240

ccg cgt gag gag tgg cgt cgc tgg cgg cgt tgg att gtg gga tac gcg     768
Pro Arg Glu Glu Trp Arg Arg Trp Arg Arg Trp Ile Val Gly Tyr Ala
                245                 250                 255

gtc tgt atg cgc ctg cat aaa aga ctt tta ttt agc cgc ttc ggt atc     816
Val Cys Met Arg Leu His Lys Arg Leu Leu Phe Ser Arg Phe Gly Ile
            260                 265                 270

ttc agt ata ttt cct atg ctg ttg gtt gtg ctt tat ggc gtt ggg att     864
Phe Ser Ile Phe Pro Met Leu Leu Val Val Leu Tyr Gly Val Gly Ile
        275                 280                 285

tat ctc act acc tgg ttt aat gaa ttc atc acc acc ggg ccg cat gga     912
Tyr Leu Thr Thr Trp Phe Asn Glu Phe Ile Thr Thr Gly Pro His Gly
    290                 295                 300

gtg gtg ttg gca atg ttt ccg ctt atc tgg gtc ggc gta gtt tgt gtt     960
Val Val Leu Ala Met Phe Pro Leu Ile Trp Val Gly Val Val Cys Val
305                 310                 315                 320

att ggt gct ttt agc gcc tgg ttt cat cgt tgc tgg ttg ttg gtg cct    1008
Ile Gly Ala Phe Ser Ala Trp Phe His Arg Cys Trp Leu Leu Val Pro
                325                 330                 335

tta gcg ccg ctt tcc gtt gtg tat gta tta tta gct tat gcc atc tgg    1056
Leu Ala Pro Leu Ser Val Val Tyr Val Leu Leu Ala Tyr Ala Ile Trp
            340                 345                 350

att att tat gga ctt att gcc ttt ttt act gga cgc gaa cct cag cgc    1104
Ile Ile Tyr Gly Leu Ile Ala Phe Phe Thr Gly Arg Glu Pro Gln Arg
        355                 360                 365

gac aaa ccc acc cgc tat tcc gca ctg gtg gaa gcg tca acc gct tat    1152
Asp Lys Pro Thr Arg Tyr Ser Ala Leu Val Glu Ala Ser Thr Ala Tyr
    370                 375                 380

tcc caa cct tct gtc aca gga act gaa aaa cta tct gaa gct taa        1197
Ser Gln Pro Ser Val Thr Gly Thr Glu Lys Leu Ser Glu Ala  *
385                 390                 395
```

<210> 238
<211> 1659
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1659)

<400> 238
```
atg att ctt gag cgc gtt gaa att gtg ggt ttt cgc ggt atc aac cgt      48
Met Ile Leu Glu Arg Val Glu Ile Val Gly Phe Arg Gly Ile Asn Arg
 1               5                  10                  15

ttg tcg ttg atg ctg gaa caa aac aac gtc ctg att ggg gag aac gcg      96
Leu Ser Leu Met Leu Glu Gln Asn Asn Val Leu Ile Gly Glu Asn Ala
            20                  25                  30

tgg ggt aaa tcc agc ttg ctg gac gcc tta act ctg ctg cta tcg cca     144
Trp Gly Lys Ser Ser Leu Leu Asp Ala Leu Thr Leu Leu Leu Ser Pro
        35                  40                  45

gaa tca gat ctc tac cat ttt gag cgc gac gat ttc tgg ttc ccg ccg     192
Glu Ser Asp Leu Tyr His Phe Glu Arg Asp Asp Phe Trp Phe Pro Pro
    50                  55                  60

gga gat atc aac ggg cga gaa cat cat ctg cat att att ttg acc ttc     240
Gly Asp Ile Asn Gly Arg Glu His His Leu His Ile Ile Leu Thr Phe
65                  70                  75                  80

cgc gaa tcg ctg cca ggc cga cat cgg gtt cgc cgt tat cgg ccg ctg     288
Arg Glu Ser Leu Pro Gly Arg His Arg Val Arg Arg Tyr Arg Pro Leu
                85                  90                  95

gaa gcg tgc tgg acg cca tgc acc gat ggc tat cac cgt att ttt tat     336
Glu Ala Cys Trp Thr Pro Cys Thr Asp Gly Tyr His Arg Ile Phe Tyr
            100                 105                 110

cgt ctg gaa ggg gag agt gcg gaa gac ggc agc gtg atg aca ctg cgc     384
Arg Leu Glu Gly Glu Ser Ala Glu Asp Gly Ser Val Met Thr Leu Arg
        115                 120                 125

agt ttt ctc gat aaa gac gga cat ccg att gat gtc gag gat att aac     432
Ser Phe Leu Asp Lys Asp Gly His Pro Ile Asp Val Glu Asp Ile Asn
    130                 135                 140

gat cag gca cgc cat ctg gtg cgt tta atg ccg gtg ctg cgc ttg cgt     480
Asp Gln Ala Arg His Leu Val Arg Leu Met Pro Val Leu Arg Leu Arg
145                 150                 155                 160

gat gcc cgt ttt atg cgc cgt att cgt aac ggc acg gtg cca aat gtc     528
Asp Ala Arg Phe Met Arg Arg Ile Arg Asn Gly Thr Val Pro Asn Val
                165                 170                 175

cct aat gtg gaa gtc acc gcg cgc cag ctc gat ttc ctc gcc cgt gag     576
Pro Asn Val Glu Val Thr Ala Arg Gln Leu Asp Phe Leu Ala Arg Glu
            180                 185                 190

tta tcc tca cat ccg caa aat ctc tct gat ggg cag att cgt cag gga     624
Leu Ser Ser His Pro Gln Asn Leu Ser Asp Gly Gln Ile Arg Gln Gly
        195                 200                 205
```

ctt tcc gca atg gta cag ctg ctt gag cat tat ttc tct gag cag ggg 672
Leu Ser Ala Met Val Gln Leu Leu Glu His Tyr Phe Ser Glu Gln Gly
210 215 220

gcc gga cag gcg cga tat cgt tta atg cgg cgg cga gcc agc aat gag 720
Ala Gly Gln Ala Arg Tyr Arg Leu Met Arg Arg Arg Ala Ser Asn Glu
225 230 235 240

caa cga agc tgg cgc tat ctg gat atc atc aac cgg atg att gac cga 768
Gln Arg Ser Trp Arg Tyr Leu Asp Ile Ile Asn Arg Met Ile Asp Arg
245 250 255

cct ggt ggg cgc tcg tat cgg gtt att ttg ctc ggc cta ttt gct act 816
Pro Gly Gly Arg Ser Tyr Arg Val Ile Leu Leu Gly Leu Phe Ala Thr
260 265 270

ttg ttg cag gca aaa ggc aca ttg cga ctg gat aaa gac gcc cgt cca 864
Leu Leu Gln Ala Lys Gly Thr Leu Arg Leu Asp Lys Asp Ala Arg Pro
275 280 285

ttg ttg ctg atc gaa gat cca gaa acc cgt tta cac ccc att atg ctt 912
Leu Leu Leu Ile Glu Asp Pro Glu Thr Arg Leu His Pro Ile Met Leu
290 295 300

tca gtt gcc tgg cat ctg ttg aat ctt ctg cca ttg cag cgc att gcc 960
Ser Val Ala Trp His Leu Leu Asn Leu Leu Pro Leu Gln Arg Ile Ala
305 310 315 320

acc acc aac tcg ggt gag ttg ctt tcg tta acg ccg gta gag cat gtt 1008
Thr Thr Asn Ser Gly Glu Leu Leu Ser Leu Thr Pro Val Glu His Val
325 330 335

tgc cga ctg gta cgt gag tcc tcg cgc gtt gcc gcc tgg cgt ctg ggg 1056
Cys Arg Leu Val Arg Glu Ser Ser Arg Val Ala Ala Trp Arg Leu Gly
340 345 350

ccg agt ggc ttg agt acc gaa gat agc cga cgc ata tcc ttt cac att 1104
Pro Ser Gly Leu Ser Thr Glu Asp Ser Arg Arg Ile Ser Phe His Ile
355 360 365

cgt ttt aac cgt ccg tca tcg ctg ttt gca cgc tgc tgg ttg ctg gtg 1152
Arg Phe Asn Arg Pro Ser Ser Leu Phe Ala Arg Cys Trp Leu Leu Val
370 375 380

gaa ggg gaa acg gaa acc tgg gtt atc aat gaa ctg gcg cgt cag tgc 1200
Glu Gly Glu Thr Glu Thr Trp Val Ile Asn Glu Leu Ala Arg Gln Cys
385 390 395 400

gga cat cat ttt gat gcc gaa ggg atc aag gtc att gag ttt gcc cag 1248
Gly His His Phe Asp Ala Glu Gly Ile Lys Val Ile Glu Phe Ala Gln
405 410 415

tcc ggg cta aag cca ctg gtt aaa ttt gcc cgc cga atg ggg att gaa 1296
Ser Gly Leu Lys Pro Leu Val Lys Phe Ala Arg Arg Met Gly Ile Glu
420 425 430

tgg cat gta ctg gtc gat ggc gat gaa gca ggg aag aaa tat gcc gct 1344
Trp His Val Leu Val Asp Gly Asp Glu Ala Gly Lys Lys Tyr Ala Ala
435 440 445

acg gta cgc agc ctg ttg aat aac gat cgg gaa gcc gaa cga gaa cat 1392
Thr Val Arg Ser Leu Leu Asn Asn Asp Arg Glu Ala Glu Arg Glu His
450 455 460

tta acg gcg tta ccg gcg ctg gat atg gaa cat ttt atg tat cgc cag 1440
Leu Thr Ala Leu Pro Ala Leu Asp Met Glu His Phe Met Tyr Arg Gln
465 470 475 480

gga ttt tcc gat gtg ttc cac cgc atg gcg caa atc ccg gaa aat gta 1488
Gly Phe Ser Asp Val Phe His Arg Met Ala Gln Ile Pro Glu Asn Val
485 490 495

ccg atg aat cta cgc aaa att atc tcg aaa gcg atc cat cgc tct tcc 1536
Pro Met Asn Leu Arg Lys Ile Ile Ser Lys Ala Ile His Arg Ser Ser
500 505 510

aaa ccc gat ctt gcc att gaa gtg gca atg gag gca gga cgt cgt ggt 1584
Lys Pro Asp Leu Ala Ile Glu Val Ala Met Glu Ala Gly Arg Arg Gly
515 520 525

gtg gac tcc gta ccg acg ctg ctg aaa aaa atg ttc tca cgc gtg ctg 1632
Val Asp Ser Val Pro Thr Leu Leu Lys Lys Met Phe Ser Arg Val Leu
530 535 540

tgg ctg gcg cgc ggt cgc gcg gat taa 1659
Trp Leu Ala Arg Gly Arg Ala Asp *
545 550

<210> 239
<211> 1059
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1059)

<400> 239
atg cta cca tct att tca atc aac aat acc agc gca gct tac cca gaa 48
Met Leu Pro Ser Ile Ser Ile Asn Asn Thr Ser Ala Ala Tyr Pro Glu
1 5 10 15

tcc atc aat gaa aat aac aat gat gaa gtt aat gga tta gta caa gag 96
Ser Ile Asn Glu Asn Asn Asn Asp Glu Val Asn Gly Leu Val Gln Glu
20 25 30

ttc aaa aac ctt ttt aat ggt aag gaa gga ata agc acc tgt att aaa 144
Phe Lys Asn Leu Phe Asn Gly Lys Glu Gly Ile Ser Thr Cys Ile Lys
35 40 45

cat cta ctt gag ctt ata aaa aac gcc ata cga gta aac gac gat cct 192
His Leu Leu Glu Leu Ile Lys Asn Ala Ile Arg Val Asn Asp Asp Pro
50 55 60

tat aga ttt aat att aat aat tcc tca gtt act tat att gat att gac 240
Tyr Arg Phe Asn Ile Asn Asn Ser Ser Val Thr Tyr Ile Asp Ile Asp
65 70 75 80

tcc aat gat aca gac cat att act att ggt atc gac aac caa gaa cca 288
Ser Asn Asp Thr Asp His Ile Thr Ile Gly Ile Asp Asn Gln Glu Pro

85 90 95

ata gaa tta cct gcg aac tat aaa gac aaa gaa ctc gtc cgt act atc 336
Ile Glu Leu Pro Ala Asn Tyr Lys Asp Lys Glu Leu Val Arg Thr Ile
100 105 110

att aat gac aac ata gtt gag aag act cat gat atc aat aac aag gaa 384
Ile Asn Asp Asn Ile Val Glu Lys Thr His Asp Ile Asn Asn Lys Glu
115 120 125

atg atc ttc agc gca tta aaa gaa ata tat gat gga gat cct ggt ttt 432
Met Ile Phe Ser Ala Leu Lys Glu Ile Tyr Asp Gly Asp Pro Gly Phe
130 135 140

atc ttc gat aag ata tca cac aaa ctc aga cat acg gta acg gaa ttt 480
Ile Phe Asp Lys Ile Ser His Lys Leu Arg His Thr Val Thr Glu Phe
145 150 155 160

gat gag agc ggg aaa agc gaa cca acg gac tta ttt acc tgg tac ggt 528
Asp Glu Ser Gly Lys Ser Glu Pro Thr Asp Leu Phe Thr Trp Tyr Gly
165 170 175

aaa gat aaa aaa ggc gac tct ctc gct att gta att aaa aat aaa aac 576
Lys Asp Lys Lys Gly Asp Ser Leu Ala Ile Val Ile Lys Asn Lys Asn
180 185 190

gga aat gat tac tta tct ctc ggt tac tac gat cag gac gac tac cac 624
Gly Asn Asp Tyr Leu Ser Leu Gly Tyr Tyr Asp Gln Asp Asp Tyr His
195 200 205

att caa aga gga att cgt att aat ggt gat agt ctc acc caa tat tgt 672
Ile Gln Arg Gly Ile Arg Ile Asn Gly Asp Ser Leu Thr Gln Tyr Cys
210 215 220

agt gaa aac gcc agg agt gct tca gcg tgg ttt gaa agc agt aaa gct 720
Ser Glu Asn Ala Arg Ser Ala Ser Ala Trp Phe Glu Ser Ser Lys Ala
225 230 235 240

atc atg gca gaa tca ttt gca act ggt tcc gat cat cag gtt gta aac 768
Ile Met Ala Glu Ser Phe Ala Thr Gly Ser Asp His Gln Val Val Asn
245 250 255

gag ctc aac ggg gaa aga ctg aga gaa cca aac gac gtt ttt aaa cgt 816
Glu Leu Asn Gly Glu Arg Leu Arg Glu Pro Asn Asp Val Phe Lys Arg
260 265 270

tat ggt cga gca ata aga tat gat ttt caa gtg gac gat gca aaa tat 864
Tyr Gly Arg Ala Ile Arg Tyr Asp Phe Gln Val Asp Asp Ala Lys Tyr
275 280 285

aaa tgc gac cat cta aaa gaa ata gtt tct act tta gtc ggt aac aaa 912
Lys Cys Asp His Leu Lys Glu Ile Val Ser Thr Leu Val Gly Asn Lys
290 295 300

att aac gtt ggc cat tct caa aaa ata tat aag cat ttt aag gat ctc 960
Ile Asn Val Gly His Ser Gln Lys Ile Tyr Lys His Phe Lys Asp Leu
305 310 315 320

gaa ggt aaa att gaa gaa agg ctt caa aat cgc cag gct gaa tat caa 1008
Glu Gly Lys Ile Glu Glu Arg Leu Gln Asn Arg Gln Ala Glu Tyr Gln
325 330 335

aat gaa att aat caa cca tct gcg cca ggt gtt aat ttt gat gat att 1056
Asn Glu Ile Asn Gln Pro Ser Ala Pro Gly Val Asn Phe Asp Asp Ile
340 345 350

taa 1059
*

<210> 240
<211> 1128
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1128)

<400> 240
atg atg cgc cat tta cgc aat att ttt aat ctg ggt atc aaa gag ttg 48
Met Met Arg His Leu Arg Asn Ile Phe Asn Leu Gly Ile Lys Glu Leu
1 5 10 15

cgc agt ctg ctc ggt gat aaa gcg atg ctg acg ctg att gtc ttc tcg 96
Arg Ser Leu Leu Gly Asp Lys Ala Met Leu Thr Leu Ile Val Phe Ser
20 25 30

ttt acg gtg tcg gtg tat tcg tca gcg acc gtt acg cca gga tcg ttg 144
Phe Thr Val Ser Val Tyr Ser Ser Ala Thr Val Thr Pro Gly Ser Leu
35 40 45

aac ctc gcg ccg atc gcc att gcc gat atg gat caa tcg cag tta tcg 192
Asn Leu Ala Pro Ile Ala Ile Ala Asp Met Asp Gln Ser Gln Leu Ser
50 55 60

aac cgg atc gtt aac agc ttc tat cgt ccg tgg ttt ttg cca ccg gag 240
Asn Arg Ile Val Asn Ser Phe Tyr Arg Pro Trp Phe Leu Pro Pro Glu
65 70 75 80

atg atc acc gcc gat gag atg gat gcc gga ctg gac gcc gga cgc tat 288
Met Ile Thr Ala Asp Glu Met Asp Ala Gly Leu Asp Ala Gly Arg Tyr
85 90 95

acc ttc gcg ata aat att ccg cct aat ttt cag cgt gat gtc ctc gcc 336
Thr Phe Ala Ile Asn Ile Pro Pro Asn Phe Gln Arg Asp Val Leu Ala
100 105 110

gga cgc cag ccg gat att cag gtg aac gtc gat gcc acg cgc atg agc 384
Gly Arg Gln Pro Asp Ile Gln Val Asn Val Asp Ala Thr Arg Met Ser
115 120 125

cag gca ttt acc ggc aat ggg tat atc cag aat att atc aac ggt gaa 432
Gln Ala Phe Thr Gly Asn Gly Tyr Ile Gln Asn Ile Ile Asn Gly Glu
130 135 140

gtg aac agc ttt gtc gcg cgc tac cgt gat aac agc gaa ccg ttg gta 480
Val Asn Ser Phe Val Ala Arg Tyr Arg Asp Asn Ser Glu Pro Leu Val
145 150 155 160

tcg ctg gaa acc cgg atg cgc ttt aac ccg aac ctc gat ccc gcg tgg 528

```
Ser Leu Glu Thr Arg Met Arg Phe Asn Pro Asn Leu Asp Pro Ala Trp
                165                 170                 175

ttt ggc ggg gtg atg gcg atc atc aac aac att acc atg ctg gcg att      576
Phe Gly Gly Val Met Ala Ile Ile Asn Asn Ile Thr Met Leu Ala Ile
            180                 185                 190

gta ttg acc gga tcg gcg ctg atc cgc gag cgt gaa cac ggc acg gtg      624
Val Leu Thr Gly Ser Ala Leu Ile Arg Glu Arg Glu His Gly Thr Val
        195                 200                 205

gaa cac tta ctg gtg atg ccg ata acg ccg ttt gag atc atg atg gcg      672
Glu His Leu Leu Val Met Pro Ile Thr Pro Phe Glu Ile Met Met Ala
    210                 215                 220

aag atc tgg tcg atg ggg ctg gtg gtg ctg gtg gta tcg gga tta tcg      720
Lys Ile Trp Ser Met Gly Leu Val Val Leu Val Val Ser Gly Leu Ser
225                 230                 235                 240

ctg gtg ctg atg gtg aaa ggt gta ctg ggc gta ccg att gaa ggc tcg      768
Leu Val Leu Met Val Lys Gly Val Leu Gly Val Pro Ile Glu Gly Ser
                245                 250                 255

atc ccg ctg ttt atg ctg ggc gtg gcg ctc agt ctg ttt gcc acc acg      816
Ile Pro Leu Phe Met Leu Gly Val Ala Leu Ser Leu Phe Ala Thr Thr
            260                 265                 270

tca atc ggc att ttt atg ggg acg ata gcg cgt tca atg ccg caa ctg      864
Ser Ile Gly Ile Phe Met Gly Thr Ile Ala Arg Ser Met Pro Gln Leu
        275                 280                 285

ggg ctg ctg gtg att ctg gtg ctg ctg ccg ctg caa atg ctt tcc ggt      912
Gly Leu Leu Val Ile Leu Val Leu Leu Pro Leu Gln Met Leu Ser Gly
    290                 295                 300

ggt tcc acg ccg cgc gaa agt atg ccg cag atg gtg cag gac att atg      960
Gly Ser Thr Pro Arg Glu Ser Met Pro Gln Met Val Gln Asp Ile Met
305                 310                 315                 320

ctg acc atg ccg acg aca cac ttt gtt agc ctc gcg cag gcc atc ctc      1008
Leu Thr Met Pro Thr Thr His Phe Val Ser Leu Ala Gln Ala Ile Leu
                325                 330                 335

tac cgg ggt gcc gga ttc gaa atc gtc tgg ccg cag ttt ctg acg ctg      1056
Tyr Arg Gly Ala Gly Phe Glu Ile Val Trp Pro Gln Phe Leu Thr Leu
            340                 345                 350

atg gca att ggc ggc gca ttt ttc acc att gcg ctg ctg cga ttc agg      1104
Met Ala Ile Gly Gly Ala Phe Phe Thr Ile Ala Leu Leu Arg Phe Arg
        355                 360                 365

aag acg att ggg aca atg gcg taa                                      1128
Lys Thr Ile Gly Thr Met Ala  *
    370                 375
```

<210> 241
<211> 2685
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(2685)

<400> 241
gtg agc cag cat tat gga aaa acc gtt gcg ctg aac aat atc act ctc 48
Met Ser Gln His Tyr Gly Lys Thr Val Ala Leu Asn Asn Ile Thr Leu
1 5 10 15

gat att ccg gcc cgc tgt atg gtc ggg ctg att ggc ccg gac ggc gtc 96
Asp Ile Pro Ala Arg Cys Met Val Gly Leu Ile Gly Pro Asp Gly Val
20 25 30

ggg aag tcg agc ttg ttg tcg ttg att tcc ggt gcc cgc gtc att gaa 144
Gly Lys Ser Ser Leu Leu Ser Leu Ile Ser Gly Ala Arg Val Ile Glu
35 40 45

cag ggc aat gtg atg gtg ctg ggc ggc gat atg cgc gac ccg aag cat 192
Gln Gly Asn Val Met Val Leu Gly Gly Asp Met Arg Asp Pro Lys His
50 55 60

cgc cgc gac gtc tgc ccg cgc atc gcc tgg atg ccg cag ggg ctg ggc 240
Arg Arg Asp Val Cys Pro Arg Ile Ala Trp Met Pro Gln Gly Leu Gly
65 70 75 80

aaa aac ctc tac cac acc ttg tcg gtg tat gaa aac gtc gat ttt ttc 288
Lys Asn Leu Tyr His Thr Leu Ser Val Tyr Glu Asn Val Asp Phe Phe
85 90 95

gct cgc ctg ttc ggt cac gac aaa gcg gag cgg gaa gtg cga atc aat 336
Ala Arg Leu Phe Gly His Asp Lys Ala Glu Arg Glu Val Arg Ile Asn
100 105 110

gag ctg ctg acc agc acc ggg tta gca ccg ttt cgc gat cgt ccg gca 384
Glu Leu Leu Thr Ser Thr Gly Leu Ala Pro Phe Arg Asp Arg Pro Ala
115 120 125

ggg aaa ctc tcc ggc ggg atg aag caa aaa ctt ggg ctg tgc tgc gcg 432
Gly Lys Leu Ser Gly Gly Met Lys Gln Lys Leu Gly Leu Cys Cys Ala
130 135 140

tta atc cac gac ccg gaa ctg ttg atc ctt gat gag cca aca acg ggg 480
Leu Ile His Asp Pro Glu Leu Leu Ile Leu Asp Glu Pro Thr Thr Gly
145 150 155 160

gtt gac ccg ctc tcc cgc tcc cag ttc tgg gat ctg atc gac agt att 528
Val Asp Pro Leu Ser Arg Ser Gln Phe Trp Asp Leu Ile Asp Ser Ile
165 170 175

cgc cag cgg cag agc aat atg agc gtg ctg gtc gcc acc gcc tat atg 576
Arg Gln Arg Gln Ser Asn Met Ser Val Leu Val Ala Thr Ala Tyr Met
180 185 190

gaa gag gcc gaa cgc ttc gac tgg ctg gta gcg atg aat gcc gga gaa 624
Glu Glu Ala Glu Arg Phe Asp Trp Leu Val Ala Met Asn Ala Gly Glu
195 200 205

gtg ctg gca act ggc agc gcc gaa gag cta cgg cag caa acg caa agc 672
Val Leu Ala Thr Gly Ser Ala Glu Glu Leu Arg Gln Gln Thr Gln Ser
210 215 220

gct acg ctg gaa gaa gca ttt ata aat ctg tta ccg caa gcg caa cgc 720
Ala Thr Leu Glu Glu Ala Phe Ile Asn Leu Leu Pro Gln Ala Gln Arg
225 230 235 240

cag gcg cat cag gcg gta gtg atc cca ccg tat caa cct gaa aac gca 768
Gln Ala His Gln Ala Val Val Ile Pro Pro Tyr Gln Pro Glu Asn Ala
245 250 255

gag att gcc atc gaa gcg cgc gat ctg acc atg cgt ttt ggt tcc ttc 816
Glu Ile Ala Ile Glu Ala Arg Asp Leu Thr Met Arg Phe Gly Ser Phe
260 265 270

gtt gcc gtt gat cac gtt aat ttc cgc att cca cgc ggg gag att ttt 864
Val Ala Val Asp His Val Asn Phe Arg Ile Pro Arg Gly Glu Ile Phe
275 280 285

ggt ttt ctt ggt tcg aac ggc tgc ggt aaa tcc acc acc atg aaa atg 912
Gly Phe Leu Gly Ser Asn Gly Cys Gly Lys Ser Thr Thr Met Lys Met
290 295 300

ctc acc gga ctg ctg ccc gcc agc gaa ggt gag gcg tgg ctg ttc ggg 960
Leu Thr Gly Leu Leu Pro Ala Ser Glu Gly Glu Ala Trp Leu Phe Gly
305 310 315 320

caa ccg gtt gat cca aaa gat atc gat acc cgc cgt cgg gtg ggc tat 1008
Gln Pro Val Asp Pro Lys Asp Ile Asp Thr Arg Arg Arg Val Gly Tyr
325 330 335

atg tcg cag gcg ttt tcg ctc tat aac gaa ctc acc gtg cgg caa aac 1056
Met Ser Gln Ala Phe Ser Leu Tyr Asn Glu Leu Thr Val Arg Gln Asn
340 345 350

ctt gag tta cat gcc cgt ttg ttt cac atc ccg gaa gcg gaa att ccc 1104
Leu Glu Leu His Ala Arg Leu Phe His Ile Pro Glu Ala Glu Ile Pro
355 360 365

gca aga gtg gct gaa atg agc gag cgt ttt aag ctc aac gac gtt gaa 1152
Ala Arg Val Ala Glu Met Ser Glu Arg Phe Lys Leu Asn Asp Val Glu
370 375 380

gat att ctg ccg gag tca ttg ccg ctc ggc att cgc cag cgg ctt tcg 1200
Asp Ile Leu Pro Glu Ser Leu Pro Leu Gly Ile Arg Gln Arg Leu Ser
385 390 395 400

ctg gcg gtg gcg gtg att cat cgc ccg gag atg tta atc ctc gat gag 1248
Leu Ala Val Ala Val Ile His Arg Pro Glu Met Leu Ile Leu Asp Glu
405 410 415

cct act tct ggt gtc gat ccg gtg gcg agg gat atg ttc tgg cag ttg 1296
Pro Thr Ser Gly Val Asp Pro Val Ala Arg Asp Met Phe Trp Gln Leu
420 425 430

atg gtc gat ctc tcg cgc cag gac aaa gtg act atc ttc atc tcc acc 1344
Met Val Asp Leu Ser Arg Gln Asp Lys Val Thr Ile Phe Ile Ser Thr
435 440 445

cac ttt atg aac gaa gcg gaa cgt tgc gac cgc atc tca ctg atg cac 1392
His Phe Met Asn Glu Ala Glu Arg Cys Asp Arg Ile Ser Leu Met His
450 455 460

gcc gga aaa gtg ctt gcc agc ggt aca ccg cag gaa ctg gtt gag aaa 1440

Ala Gly Lys Val Leu Ala Ser Gly Thr Pro Gln Glu Leu Val Glu Lys
465 470 475 480

cgc gga gcc gcc agt ctg gaa gag gca ttt atc gcc tat ttg cag gaa 1488
Arg Gly Ala Ala Ser Leu Glu Glu Ala Phe Ile Ala Tyr Leu Gln Glu
485 490 495

gcg gca ggg cag agc aac gaa gcc gaa gcg ccg ccc gtg gta cac gac 1536
Ala Ala Gly Gln Ser Asn Glu Ala Glu Ala Pro Pro Val Val His Asp
500 505 510

acc acc cac gcg ccg cgt cag gga ttt agc ctg cgc cgt ctg ttt agc 1584
Thr Thr His Ala Pro Arg Gln Gly Phe Ser Leu Arg Arg Leu Phe Ser
515 520 525

tac agc cgc cgc gaa gcg ctg gaa ctg cga cgc gat cca gta cgt tcg 1632
Tyr Ser Arg Arg Glu Ala Leu Glu Leu Arg Arg Asp Pro Val Arg Ser
530 535 540

acg ctg gcg ctg atg gga acg gtg atc ctg atg ctg ata atg ggt tac 1680
Thr Leu Ala Leu Met Gly Thr Val Ile Leu Met Leu Ile Met Gly Tyr
545 550 555 560

ggc atc agt atg gat gtg gaa aac ctg cgc ttt gcg gtg ctc gac cgc 1728
Gly Ile Ser Met Asp Val Glu Asn Leu Arg Phe Ala Val Leu Asp Arg
565 570 575

gac cag acc gtc agt agc cag gcg tgg aca ctc aac ctc tcc ggt tcc 1776
Asp Gln Thr Val Ser Ser Gln Ala Trp Thr Leu Asn Leu Ser Gly Ser
580 585 590

cgt tac ttt atc gaa cag ccg ccg ctc acc agt tat gac gag ctt gat 1824
Arg Tyr Phe Ile Glu Gln Pro Pro Leu Thr Ser Tyr Asp Glu Leu Asp
595 600 605

cgt cgg atg cgt gcg ggc gat atc acg gtg gcg att gag atc ccg ccc 1872
Arg Arg Met Arg Ala Gly Asp Ile Thr Val Ala Ile Glu Ile Pro Pro
610 615 620

aat ttc ggg cgc gat atc gcg cgt ggt acg cct gtg gaa ctc ggc gtc 1920
Asn Phe Gly Arg Asp Ile Ala Arg Gly Thr Pro Val Glu Leu Gly Val
625 630 635 640

tgg atc gac gga gcg atg ccg agc cgt gct gaa acg gta aaa ggt tac 1968
Trp Ile Asp Gly Ala Met Pro Ser Arg Ala Glu Thr Val Lys Gly Tyr
645 650 655

gtg cag gcc atg cac cag agc tgg tta cag gat gtg gcg agc cga caa 2016
Val Gln Ala Met His Gln Ser Trp Leu Gln Asp Val Ala Ser Arg Gln
660 665 670

tcg aca ccc gcc agc caa agc ggg ctg atg aat att gag acg cgc tat 2064
Ser Thr Pro Ala Ser Gln Ser Gly Leu Met Asn Ile Glu Thr Arg Tyr
675 680 685

cgc tat aac ccg gac gta aaa agc ctg cca gcg att gtt ccg gcg gtg 2112
Arg Tyr Asn Pro Asp Val Lys Ser Leu Pro Ala Ile Val Pro Ala Val
690 695 700

atc ccg ctt ctg ctg atg atg atc ccg tca atg cta agc gcc ctt agc 2160
Ile Pro Leu Leu Leu Met Met Ile Pro Ser Met Leu Ser Ala Leu Ser

705 710 715 720

gtg gtg cgg gaa aaa gag ctt ggg tcg att atc aac ctt tac gtg acc 2208
Val Val Arg Glu Lys Glu Leu Gly Ser Ile Ile Asn Leu Tyr Val Thr
725 730 735

ccc acc acg cgt agt gaa ttt ttg ctt ggt aaa cag ttg cca tac atc 2256
Pro Thr Thr Arg Ser Glu Phe Leu Leu Gly Lys Gln Leu Pro Tyr Ile
740 745 750

gcg ctg ggg atg ctg aac ttt ttc ctg ctc tgc ggc ctg tcg gtg ttt 2304
Ala Leu Gly Met Leu Asn Phe Phe Leu Leu Cys Gly Leu Ser Val Phe
755 760 765

gtg ttt ggc gta ccg cat aaa ggc agt ttc ctg acg ctc acc ctg gcg 2352
Val Phe Gly Val Pro His Lys Gly Ser Phe Leu Thr Leu Thr Leu Ala
770 775 780

gcg ctg ctg tat atc atc att gcc acc gga atg ggg ctg ctg atc tcc 2400
Ala Leu Leu Tyr Ile Ile Ile Ala Thr Gly Met Gly Leu Leu Ile Ser
785 790 795 800

acc ttt atg aaa agc cag att gcc gcc att ttc gga acg gcg att atc 2448
Thr Phe Met Lys Ser Gln Ile Ala Ala Ile Phe Gly Thr Ala Ile Ile
805 810 815

acg ttg atc ccg gcg aca cag ttt tcc ggg atg atc gat ccg gta gct 2496
Thr Leu Ile Pro Ala Thr Gln Phe Ser Gly Met Ile Asp Pro Val Ala
820 825 830

tcg ctg gaa ggg cct gga cgt tgg atc ggc gag gtt tac ccg acc agt 2544
Ser Leu Glu Gly Pro Gly Arg Trp Ile Gly Glu Val Tyr Pro Thr Ser
835 840 845

cat ttt ctg act atc gcc cgc ggg acg ttc tcg aaa gcg ctg gat ctg 2592
His Phe Leu Thr Ile Ala Arg Gly Thr Phe Ser Lys Ala Leu Asp Leu
850 855 860

act gat ttg tgg caa ctt ttt atc ccg tta ctg ata gcc atc ccg ctg 2640
Thr Asp Leu Trp Gln Leu Phe Ile Pro Leu Leu Ile Ala Ile Pro Leu
865 870 875 880

gtg atg ggc tta agt atc ctg ctg ctg aaa aaa cag gag gga tga 2685
Val Met Gly Leu Ser Ile Leu Leu Leu Lys Lys Gln Glu Gly *
885 890

<210> 242
<211> 1068
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1068)

<400> 242
atg gat aag agt aag cgc cat ctg gcg tgg tgg gtt gtc ggg tta ctg 48
Met Asp Lys Ser Lys Arg His Leu Ala Trp Trp Val Val Gly Leu Leu
1 5 10 15

gcg gtg gcg gct atc gtg gcg tgg tgg ctg ttg cgc ccg gca ggt gtg 96
Ala Val Ala Ala Ile Val Ala Trp Trp Leu Leu Arg Pro Ala Gly Val
20 25 30

ccg gaa ggc ttt gct gtc agt aat ggg cgc att gaa gcg acg gaa gtg 144
Pro Glu Gly Phe Ala Val Ser Asn Gly Arg Ile Glu Ala Thr Glu Val
35 40 45

gat att gcc agc aaa att gcc ggg cgt atc gac acc att ctg gtg aaa 192
Asp Ile Ala Ser Lys Ile Ala Gly Arg Ile Asp Thr Ile Leu Val Lys
50 55 60

gaa ggc aag ttt gtt cgc gaa ggt gaa gtg ctg gcg aag atg gat act 240
Glu Gly Lys Phe Val Arg Glu Gly Glu Val Leu Ala Lys Met Asp Thr
65 70 75 80

cgc gtg ttg cag gaa cag cga ctg gaa gcc atc gcg caa atc aaa gag 288
Arg Val Leu Gln Glu Gln Arg Leu Glu Ala Ile Ala Gln Ile Lys Glu
85 90 95

gca caa agc gcc gtt gct gcc gcg cag gct ttg ctg gag caa cga caa 336
Ala Gln Ser Ala Val Ala Ala Ala Gln Ala Leu Leu Glu Gln Arg Gln
100 105 110

agc gaa act cgt gcc gca cag tcg ctg gtt aat caa cgc cag gca gaa 384
Ser Glu Thr Arg Ala Ala Gln Ser Leu Val Asn Gln Arg Gln Ala Glu
115 120 125

ctg gac tcc gta gca aaa cgt cat acg cgt tcc cgt tca ctg gcc caa 432
Leu Asp Ser Val Ala Lys Arg His Thr Arg Ser Arg Ser Leu Ala Gln
130 135 140

cga ggg gct att tct gcg caa cag ctg gat gac gat cgc gcc gcc gct 480
Arg Gly Ala Ile Ser Ala Gln Gln Leu Asp Asp Asp Arg Ala Ala Ala
145 150 155 160

gag agc gcc cga gct gcg ctg gaa tcg gcg aaa gct cag gta tcg gct 528
Glu Ser Ala Arg Ala Ala Leu Glu Ser Ala Lys Ala Gln Val Ser Ala
165 170 175

tct aaa gcg gct ata gaa gcg gca cgc acc aat atc att cag gcg caa 576
Ser Lys Ala Ala Ile Glu Ala Ala Arg Thr Asn Ile Ile Gln Ala Gln
180 185 190

acc cgc gtc gaa gcg gca caa gcc act gaa cgg cgc att gcc gca gat 624
Thr Arg Val Glu Ala Ala Gln Ala Thr Glu Arg Arg Ile Ala Ala Asp
195 200 205

atc gat gac agc gaa ctg aaa gcc ccg cgt gac gga cgc gtg cag tat 672
Ile Asp Asp Ser Glu Leu Lys Ala Pro Arg Asp Gly Arg Val Gln Tyr
210 215 220

cgg gtt gcc gag cca ggc gaa gtg ctg gcg gca ggc ggt cgg gtg ctg 720
Arg Val Ala Glu Pro Gly Glu Val Leu Ala Ala Gly Gly Arg Val Leu
225 230 235 240

aat atg gtc gat ctc agc gac gtc tat atg act ttc ttc ctg cca acc 768
Asn Met Val Asp Leu Ser Asp Val Tyr Met Thr Phe Phe Leu Pro Thr
245 250 255

gaa cag gcg ggc acg ctg aaa ctg ggc ggt gaa gcc cgg ctg atc ctc 816
Glu Gln Ala Gly Thr Leu Lys Leu Gly Gly Glu Ala Arg Leu Ile Leu
260 265 270

gat gcc gcg cca gat ctg cgt att cct gca acc atc agt ttt gtc gcc 864
Asp Ala Ala Pro Asp Leu Arg Ile Pro Ala Thr Ile Ser Phe Val Ala
275 280 285

agt gtc gcc cag ttc acg cca aaa acc gtc gaa acc agc gat gaa cgg 912
Ser Val Ala Gln Phe Thr Pro Lys Thr Val Glu Thr Ser Asp Glu Arg
290 295 300

ctg aaa ctg atg ttc cgc gtc aaa gcg cgt atc cca ccg gaa tta ctc 960
Leu Lys Leu Met Phe Arg Val Lys Ala Arg Ile Pro Pro Glu Leu Leu
305 310 315 320

cag cag cat ctg gaa tat gtc aaa acc ggt ttg ccg ggc gta gcg tgg 1008
Gln Gln His Leu Glu Tyr Val Lys Thr Gly Leu Pro Gly Val Ala Trp
325 330 335

gtg cgg gtg aat gaa gaa ctt ccg tgg cct gac gac ctc gtg gtg agg 1056
Val Arg Val Asn Glu Glu Leu Pro Trp Pro Asp Asp Leu Val Val Arg
340 345 350

ttg ccg caa tga 1068
Leu Pro Gln *
355

<210> 243
<211> 708
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(708)

<400> 243
atg tac cgt tat ttg tct att gct gcg gtg gta ctg agc gca gca ttt 48
Met Tyr Arg Tyr Leu Ser Ile Ala Ala Val Val Leu Ser Ala Ala Phe
1 5 10 15

tcc ggc ccg gcg ttg gcc gaa ggt atc aat agt ttt tct cag gcg aaa 96
Ser Gly Pro Ala Leu Ala Glu Gly Ile Asn Ser Phe Ser Gln Ala Lys
20 25 30

gcc gcg gcg gta aaa gtc cac gct gac gcg ccc ggt acg ttt tat tgc 144
Ala Ala Ala Val Lys Val His Ala Asp Ala Pro Gly Thr Phe Tyr Cys
35 40 45

gga tgt aaa att aac tgg cag ggc aaa aaa ggc gtt gtt gat ctg caa 192
Gly Cys Lys Ile Asn Trp Gln Gly Lys Lys Gly Val Val Asp Leu Gln
50 55 60

tcg tgc ggc tat cag gtg cgc aaa aat gaa aac cgc gcc agc cgc gta 240
Ser Cys Gly Tyr Gln Val Arg Lys Asn Glu Asn Arg Ala Ser Arg Val
65 70 75 80

gag tgg gaa cat gtc gtt ccc gcc tgg cag ttc ggt cac cag cgc cag 288
Glu Trp Glu His Val Val Pro Ala Trp Gln Phe Gly His Gln Arg Gln

```
                85                  90                  95

tgc tgg cag gac ggt gga cgt aaa aac tgc gct aaa gat ccg gtc tat     336
Cys Trp Gln Asp Gly Gly Arg Lys Asn Cys Ala Lys Asp Pro Val Tyr
            100                 105                 110

cgc aag atg gaa agc gat atg cat aac ctg cag ccg tca gtc ggt gag     384
Arg Lys Met Glu Ser Asp Met His Asn Leu Gln Pro Ser Val Gly Glu
        115                 120                 125

gtg aat ggc gat cgc ggc aac ttt atg tac agc cag tgg aat ggc ggt     432
Val Asn Gly Asp Arg Gly Asn Phe Met Tyr Ser Gln Trp Asn Gly Gly
    130                 135                 140

gaa ggc cag tac ggt caa tgc gcc atg aag gtc gat ttc aaa gaa aaa     480
Glu Gly Gln Tyr Gly Gln Cys Ala Met Lys Val Asp Phe Lys Glu Lys
145                 150                 155                 160

gct gcc gaa cca cca gcg cgt gca cgc ggt gcc att gcg cgc acc tac     528
Ala Ala Glu Pro Pro Ala Arg Ala Arg Gly Ala Ile Ala Arg Thr Tyr
                165                 170                 175

ttc tat atg cgc gac caa tac aac ctg aca ctc tct cgc cag caa acg     576
Phe Tyr Met Arg Asp Gln Tyr Asn Leu Thr Leu Ser Arg Gln Gln Thr
            180                 185                 190

cag ctg ttc aac gca tgg aac aag atg tat ccg gtt acc gac tgg gag     624
Gln Leu Phe Asn Ala Trp Asn Lys Met Tyr Pro Val Thr Asp Trp Glu
        195                 200                 205

tgc gag cgc gat gaa cgc atc gcg aag gtg cag ggc aat cat aac ccg     672
Cys Glu Arg Asp Glu Arg Ile Ala Lys Val Gln Gly Asn His Asn Pro
    210                 215                 220

tat gtg caa cgc gct tgc cag gcg cga aag agc taa                     708
Tyr Val Gln Arg Ala Cys Gln Ala Arg Lys Ser  *
225                 230                 235
```

<210> 244
<211> 1443
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1443)

<400> 244

```
atg tcc aga agg ctt cgc aga aca aaa atc gtt acc acg tta ggc cca     48
Met Ser Arg Arg Leu Arg Arg Thr Lys Ile Val Thr Thr Leu Gly Pro
 1               5                  10                  15

gca aca gat cgc gat aat aat ctt gaa aaa gtt atc gcg gcg ggt gcc     96
Ala Thr Asp Arg Asp Asn Asn Leu Glu Lys Val Ile Ala Ala Gly Ala
            20                  25                  30

aac gtt gta cgt atg aac ttt tct cac ggc tcg cct gaa gat cac aaa     144
Asn Val Val Arg Met Asn Phe Ser His Gly Ser Pro Glu Asp His Lys
        35                  40                  45
```

```
atg cgc gcg gat aaa gtt cgt gag att gcc gca aaa ctg ggg cgt cat    192
Met Arg Ala Asp Lys Val Arg Glu Ile Ala Ala Lys Leu Gly Arg His
     50                  55                  60

gtg gct att ctg ggt gac ctc cag ggg ccc aaa atc cgt gta tcc acc    240
Val Ala Ile Leu Gly Asp Leu Gln Gly Pro Lys Ile Arg Val Ser Thr
 65                  70                  75                  80

ttt aaa gaa ggc aaa gtt ttc ctc aat att ggg gat aaa ttc ctg ctc    288
Phe Lys Glu Gly Lys Val Phe Leu Asn Ile Gly Asp Lys Phe Leu Leu
                 85                  90                  95

gac gcc aac ctg ggt aaa ggt gaa ggc gac aaa gaa aaa gtc ggt atc    336
Asp Ala Asn Leu Gly Lys Gly Glu Gly Asp Lys Glu Lys Val Gly Ile
            100                 105                 110

gac tac aaa ggc ctg cct gct gac gtc gtg cct ggt gac atc ctg ctg    384
Asp Tyr Lys Gly Leu Pro Ala Asp Val Val Pro Gly Asp Ile Leu Leu
        115                 120                 125

ctg gac gat ggt cgc gtc cag tta aaa gta ctg gaa gtt cag ggc atg    432
Leu Asp Asp Gly Arg Val Gln Leu Lys Val Leu Glu Val Gln Gly Met
    130                 135                 140

aaa gtg ttc acc gaa gtc acc gtc ggt ggt ccc ctc tcc aac aat aaa    480
Lys Val Phe Thr Glu Val Thr Val Gly Gly Pro Leu Ser Asn Asn Lys
145                 150                 155                 160

ggt atc aac aaa ctt ggc ggc ggt ttg tcg gct gaa gcg ctg acc gaa    528
Gly Ile Asn Lys Leu Gly Gly Gly Leu Ser Ala Glu Ala Leu Thr Glu
                165                 170                 175

aaa gac aaa gca gac att aag act gcg gcg ttg att ggc gta gat tac    576
Lys Asp Lys Ala Asp Ile Lys Thr Ala Ala Leu Ile Gly Val Asp Tyr
            180                 185                 190

ctg gct gtc tcc ttc cca cgc tgt ggc gaa gat ctg aac tat gcc cgt    624
Leu Ala Val Ser Phe Pro Arg Cys Gly Glu Asp Leu Asn Tyr Ala Arg
        195                 200                 205

cgc ctg gca cgc gat gca gga tgt gat gcg aaa att gtt gcc aag gtt    672
Arg Leu Ala Arg Asp Ala Gly Cys Asp Ala Lys Ile Val Ala Lys Val
    210                 215                 220

gaa cgt gcg gaa gcc gtt tgc agc cag gat gca atg gat gac atc atc    720
Glu Arg Ala Glu Ala Val Cys Ser Gln Asp Ala Met Asp Asp Ile Ile
225                 230                 235                 240

ctc gcc tct gac gtg gta atg gtt gca cgt ggc gac ctc ggt gtg gaa    768
Leu Ala Ser Asp Val Val Met Val Ala Arg Gly Asp Leu Gly Val Glu
                245                 250                 255

att ggc gac ccg gaa ctg gtc ggc att cag aaa gcg ttg atc cgt cgt    816
Ile Gly Asp Pro Glu Leu Val Gly Ile Gln Lys Ala Leu Ile Arg Arg
            260                 265                 270

gcg cgt cag cta aac cga gcg gta atc acg gcg acc cag atg atg gag    864
Ala Arg Gln Leu Asn Arg Ala Val Ile Thr Ala Thr Gln Met Met Glu
        275                 280                 285

tca atg att act aac ccg atg ccg acg cgt gca gaa gtc atg gac gta    912
```

```
Ser Met Ile Thr Asn Pro Met Pro Thr Arg Ala Glu Val Met Asp Val
    290                 295                 300

gca aac gcc gtt ctg gat ggt act gac gct gtg atg ctg tct gca gaa    960
Ala Asn Ala Val Leu Asp Gly Thr Asp Ala Val Met Leu Ser Ala Glu
305                 310                 315                 320

act gcc gct ggg cag tat ccg tca gaa acc gtt gca gcc atg gcg cgc    1008
Thr Ala Ala Gly Gln Tyr Pro Ser Glu Thr Val Ala Ala Met Ala Arg
                325                 330                 335

gtt tgc ctg ggt gcg gaa aaa atc ccg agc atc aac gtt tct aaa cac    1056
Val Cys Leu Gly Ala Glu Lys Ile Pro Ser Ile Asn Val Ser Lys His
            340                 345                 350

cgt ctg gac gtt cag ttc gac aat gtg gaa gaa gct att gcc atg tca    1104
Arg Leu Asp Val Gln Phe Asp Asn Val Glu Glu Ala Ile Ala Met Ser
        355                 360                 365

gca atg tac gca gct aac cac ctg aaa ggc gtt acg gcg atc atc acc    1152
Ala Met Tyr Ala Ala Asn His Leu Lys Gly Val Thr Ala Ile Ile Thr
    370                 375                 380

atg acc gaa tcg ggt cgt acc gcg ctg atg acc tcc cgt atc agc tct    1200
Met Thr Glu Ser Gly Arg Thr Ala Leu Met Thr Ser Arg Ile Ser Ser
385                 390                 395                 400

ggt ctg cca att ttc gcc atg tcg cgc cat gaa cgt acg ctg aac ctg    1248
Gly Leu Pro Ile Phe Ala Met Ser Arg His Glu Arg Thr Leu Asn Leu
                405                 410                 415

act gct ctc tat cgt ggc gtt acg ccg gtg cac ttt gat agc gct aat    1296
Thr Ala Leu Tyr Arg Gly Val Thr Pro Val His Phe Asp Ser Ala Asn
            420                 425                 430

gac ggc gta gca gct gcc agc gaa gcg gtt aat ctg ctg cgc gat aaa    1344
Asp Gly Val Ala Ala Ala Ser Glu Ala Val Asn Leu Leu Arg Asp Lys
        435                 440                 445

ggt tac ttg atg tct ggt gac ctg gtg att gtc acc cag ggc gac gtg    1392
Gly Tyr Leu Met Ser Gly Asp Leu Val Ile Val Thr Gln Gly Asp Val
    450                 455                 460

atg agt acc gtg ggt tct act aat acc acg cgt att tta acg gta gag    1440
Met Ser Thr Val Gly Ser Thr Asn Thr Thr Arg Ile Leu Thr Val Glu
465                 470                 475                 480

taa                                                                1443
 *
```

<210> 245
<211> 720
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(720)

<400> 245

```
atg atc aat gta tta att atc gat gac gac gca atg gtc gcg gag ctg      48
Met Ile Asn Val Leu Ile Ile Asp Asp Asp Ala Met Val Ala Glu Leu
 1               5                  10                  15

aat cgc cga tac gta gca caa atc cca ggc ttt caa tgc tgt gga aca      96
Asn Arg Arg Tyr Val Ala Gln Ile Pro Gly Phe Gln Cys Cys Gly Thr
            20                  25                  30

gcc tcg acg ctg gag aaa gcc aaa gag att atc ttc aat agc gat acg     144
Ala Ser Thr Leu Glu Lys Ala Lys Glu Ile Ile Phe Asn Ser Asp Thr
        35                  40                  45

cct atc gac ctg ata ttg ctc gat atc tat atg caa aaa gag aac ggg     192
Pro Ile Asp Leu Ile Leu Leu Asp Ile Tyr Met Gln Lys Glu Asn Gly
    50                  55                  60

ctc gat tta ctg cct gtc ctg cat aac gcg cgt tgc aaa agt gat gtg     240
Leu Asp Leu Leu Pro Val Leu His Asn Ala Arg Cys Lys Ser Asp Val
65                  70                  75                  80

att gtc atc tcc tcc gca gcc gat gcg gca acc att aaa gat tcg ctg     288
Ile Val Ile Ser Ser Ala Ala Asp Ala Ala Thr Ile Lys Asp Ser Leu
                85                  90                  95

cat tac ggt gtc gtg gat tac ctg atc aaa ccc ttc cag gct tcc cgc     336
His Tyr Gly Val Val Asp Tyr Leu Ile Lys Pro Phe Gln Ala Ser Arg
            100                 105                 110

ttt gaa gag gcg ctc acc ggc tgg cgg caa aag aaa atg gcg ctg gaa     384
Phe Glu Glu Ala Leu Thr Gly Trp Arg Gln Lys Lys Met Ala Leu Glu
        115                 120                 125

aaa cat cag tat tat gac cag gcc gag ctt gat cag cta atc cac ggc     432
Lys His Gln Tyr Tyr Asp Gln Ala Glu Leu Asp Gln Leu Ile His Gly
    130                 135                 140

agc agc tcc aac gaa cag gat cct cgt cgc ttg cca aaa ggc tta acg     480
Ser Ser Ser Asn Glu Gln Asp Pro Arg Arg Leu Pro Lys Gly Leu Thr
145                 150                 155                 160

ccg cag acg tta cgc aca ctg tgc cag tgg att gac gcg cat cag gac     528
Pro Gln Thr Leu Arg Thr Leu Cys Gln Trp Ile Asp Ala His Gln Asp
                165                 170                 175

tat gaa ttt tca acc gac gag cta gcc aac gag gtt aac att tcg cgt     576
Tyr Glu Phe Ser Thr Asp Glu Leu Ala Asn Glu Val Asn Ile Ser Arg
            180                 185                 190

gtt tcc tgc cgt aaa tac ctc atc tgg ctg gtc aac tgc cac atc ttg     624
Val Ser Cys Arg Lys Tyr Leu Ile Trp Leu Val Asn Cys His Ile Leu
        195                 200                 205

ttc acc agt atc cat tat ggc gtc acg ggg cgt cca gtt tac cgt tac     672
Phe Thr Ser Ile His Tyr Gly Val Thr Gly Arg Pro Val Tyr Arg Tyr
    210                 215                 220

cgt att cag gca gaa cac tac tca cta ctg aaa caa tat tgc caa taa     720
Arg Ile Gln Ala Glu His Tyr Ser Leu Leu Lys Gln Tyr Cys Gln  *
225                 230                 235
```

<210> 246
<211> 1632
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1632)

<400> 246
```
atg aga cat tca ttg ccc tac cgc atg tta cgc aaa cgt ccg atg aaa      48
Met Arg His Ser Leu Pro Tyr Arg Met Leu Arg Lys Arg Pro Met Lys
 1               5                  10                  15

ttg agt acc aca gtg atc tta atg gtc agt gcg gta ctg ttc tcg gtg      96
Leu Ser Thr Thr Val Ile Leu Met Val Ser Ala Val Leu Phe Ser Val
            20                  25                  30

cta ttg gtg gtg cat ctg att tac ttc tcg caa atc agt gat atg acg     144
Leu Leu Val Val His Leu Ile Tyr Phe Ser Gln Ile Ser Asp Met Thr
        35                  40                  45

cga gat ggg cta gcc aac aag gca ctg gca gtg gcg cgt acc ctc gcc     192
Arg Asp Gly Leu Ala Asn Lys Ala Leu Ala Val Ala Arg Thr Leu Ala
    50                  55                  60

gac tcg ccg gaa atc cgt cag ggc ttg cag aaa aaa ccg cag gag agt     240
Asp Ser Pro Glu Ile Arg Gln Gly Leu Gln Lys Lys Pro Gln Glu Ser
 65                  70                  75                  80

ggc atc cag gcc atc gcg gaa gcc gta cgc aaa cgc aac gat ctg ctg     288
Gly Ile Gln Ala Ile Ala Glu Ala Val Arg Lys Arg Asn Asp Leu Leu
                 85                  90                  95

ttt att gtc gtt acc gat atg caa agt ctt cgc tac tcg cat cct gaa     336
Phe Ile Val Val Thr Asp Met Gln Ser Leu Arg Tyr Ser His Pro Glu
            100                 105                 110

gcc cag cgt att ggt cag cca ttt aaa ggt gat gac atc ctt aaa gcg     384
Ala Gln Arg Ile Gly Gln Pro Phe Lys Gly Asp Asp Ile Leu Lys Ala
        115                 120                 125

ctg aat ggc gaa gaa aat gtc gct atc aat cgc ggt ttt ctg gcg cag     432
Leu Asn Gly Glu Glu Asn Val Ala Ile Asn Arg Gly Phe Leu Ala Gln
    130                 135                 140

gct tta cgc gta ttt acc ccc atc tac gat gaa aat cat aaa caa att     480
Ala Leu Arg Val Phe Thr Pro Ile Tyr Asp Glu Asn His Lys Gln Ile
145                 150                 155                 160

ggc gtg gtg gcg atc ggc ctt gag tta agc cgt gtg acc caa cag atc     528
Gly Val Val Ala Ile Gly Leu Glu Leu Ser Arg Val Thr Gln Gln Ile
                165                 170                 175

aat gac agt cgc tgg agc att atc tgg tcg gta tta ttt ggc atg ctg     576
Asn Asp Ser Arg Trp Ser Ile Ile Trp Ser Val Leu Phe Gly Met Leu
            180                 185                 190

gtc gga ctg att ggc acc tgc att ctg gtt aag gta ctg aaa aaa atc     624
```

Val Gly Leu Ile Gly Thr Cys Ile Leu Val Lys Val Leu Lys Lys Ile
195 200 205

ctt ttc ggc ctg gaa ccc tac gaa atc tcc acg ctg ttt gag caa cgc 672
Leu Phe Gly Leu Glu Pro Tyr Glu Ile Ser Thr Leu Phe Glu Gln Arg
210 215 220

cag gcc atg ttg cag tct atc aaa gaa ggc gtc gtt gcc gtg gac gat 720
Gln Ala Met Leu Gln Ser Ile Lys Glu Gly Val Val Ala Val Asp Asp
225 230 235 240

cgc ggc gag gtc acg ctg atc aac gat gcc gca caa gaa ttg ctg aat 768
Arg Gly Glu Val Thr Leu Ile Asn Asp Ala Ala Gln Glu Leu Leu Asn
245 250 255

tac cgt aag tcg cag gac gat gag aaa ctg tcg acg cta agc cac tca 816
Tyr Arg Lys Ser Gln Asp Asp Glu Lys Leu Ser Thr Leu Ser His Ser
260 265 270

tgg tca cag gtg gta gat gtc tcg gaa gtg tta cgc gac ggt acc ccg 864
Trp Ser Gln Val Val Asp Val Ser Glu Val Leu Arg Asp Gly Thr Pro
275 280 285

cgc cgc gac gaa gag att acg att aaa gac cgg cta tta ctg atc aac 912
Arg Arg Asp Glu Glu Ile Thr Ile Lys Asp Arg Leu Leu Leu Ile Asn
290 295 300

acc gtt ccg gtg cgc agt aat ggc gtt atc atc ggt gcc att tca acc 960
Thr Val Pro Val Arg Ser Asn Gly Val Ile Ile Gly Ala Ile Ser Thr
305 310 315 320

ttc agg gac aaa act gaa gta cgt aaa ctg atg cag cga ctc gac ggt 1008
Phe Arg Asp Lys Thr Glu Val Arg Lys Leu Met Gln Arg Leu Asp Gly
325 330 335

ctg gtc aac tat gct gac gca ctt cgt gaa cga tcc cac gaa ttt atg 1056
Leu Val Asn Tyr Ala Asp Ala Leu Arg Glu Arg Ser His Glu Phe Met
340 345 350

aat aaa ttg cat gtg att ctc gga tta ttg cat ctg aag agt tat aag 1104
Asn Lys Leu His Val Ile Leu Gly Leu Leu His Leu Lys Ser Tyr Lys
355 360 365

cag ttg gaa gat tac att ctc aaa aca gcc aat aac tat cag gaa gag 1152
Gln Leu Glu Asp Tyr Ile Leu Lys Thr Ala Asn Asn Tyr Gln Glu Glu
370 375 380

att ggc tct ctg ctg ggt aag atc aaa tct ccg gtt atc gct ggt ttt 1200
Ile Gly Ser Leu Leu Gly Lys Ile Lys Ser Pro Val Ile Ala Gly Phe
385 390 395 400

tta atc agc aag att aac cgc gcg acc gat tta ggc cat acg ctg att 1248
Leu Ile Ser Lys Ile Asn Arg Ala Thr Asp Leu Gly His Thr Leu Ile
405 410 415

tta aac agt gaa agc cag ctg cca gac agc ggc agt gag gac cag gtc 1296
Leu Asn Ser Glu Ser Gln Leu Pro Asp Ser Gly Ser Glu Asp Gln Val
420 425 430

gcg acg ctg att acc acg ttg gga aat ctg ata gaa aac gcg ctg gag 1344
Ala Thr Leu Ile Thr Thr Leu Gly Asn Leu Ile Glu Asn Ala Leu Glu

```
        435                 440                 445

gca tta ggg ccg gaa ccc gga ggc gaa att agc gta aca ttg cac tac    1392
Ala Leu Gly Pro Glu Pro Gly Gly Glu Ile Ser Val Thr Leu His Tyr
    450                 455                 460

cgt cac ggc tgg ctg cac tgt gaa gtt aat gat gat gga ccg ggg atc    1440
Arg His Gly Trp Leu His Cys Glu Val Asn Asp Asp Gly Pro Gly Ile
465                 470                 475                 480

gca ccc gat aaa atc gat cac att ttt gac aaa ggt gtc tcg aca aaa    1488
Ala Pro Asp Lys Ile Asp His Ile Phe Asp Lys Gly Val Ser Thr Lys
                485                 490                 495

gga agc gag cga ggc gtc ggt tta gca ctt gtc aaa caa cag gta gaa    1536
Gly Ser Glu Arg Gly Val Gly Leu Ala Leu Val Lys Gln Gln Val Glu
            500                 505                 510

aat ctc ggc ggc agc atc gcc gtg gaa tcg gaa ccc ggg att ttc aca    1584
Asn Leu Gly Gly Ser Ile Ala Val Glu Ser Glu Pro Gly Ile Phe Thr
        515                 520                 525

caa ttt ttt gtc cag ata ccc tgg gac ggg gag agg tcg aac aga tga    1632
Gln Phe Phe Val Gln Ile Pro Trp Asp Gly Glu Arg Ser Asn Arg  *
    530                 535                 540
```

<210> 247
<211> 987
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(987)

<400> 247

```
ttg agt gta ccg ctg tcg aca tgg aat ctt ctg cga tac aac aat tcg    48
Met Ser Val Pro Leu Ser Thr Trp Asn Leu Leu Arg Tyr Asn Asn Ser
 1               5                  10                  15

tat cta cag aag gta act atg ttt cca caa tgc aaa ttt tcc cgc gag    96
Tyr Leu Gln Lys Val Thr Met Phe Pro Gln Cys Lys Phe Ser Arg Glu
            20                  25                  30

ttt cta cat cct cgc tac tgg ctc aca tgg ttt ggg ctt ggt gta ctc    144
Phe Leu His Pro Arg Tyr Trp Leu Thr Trp Phe Gly Leu Gly Val Leu
        35                  40                  45

tgg ctt tgg gta cag ctt cct tat cct gtt ctc tgc ttt ctc ggc acg    192
Trp Leu Trp Val Gln Leu Pro Tyr Pro Val Leu Cys Phe Leu Gly Thr
    50                  55                  60

cgt att ggc gca atg gcg cga cca ttc ctg aaa cgt cgt gaa tct atc    240
Arg Ile Gly Ala Met Ala Arg Pro Phe Leu Lys Arg Arg Glu Ser Ile
65                  70                  75                  80

gcc cgt aaa aac ctg gaa ctt tgt ttc ccg cag cat tct gcg gaa gaa    288
Ala Arg Lys Asn Leu Glu Leu Cys Phe Pro Gln His Ser Ala Glu Glu
                85                  90                  95
```

```
cgc gag aag atg att gcc gaa aac ttt cgt tca ctc ggc atg gcg ctg     336
Arg Glu Lys Met Ile Ala Glu Asn Phe Arg Ser Leu Gly Met Ala Leu
            100                 105                 110

gta gaa acc ggc atg gca tgg ttc tgg ccc gac agt cgc gta cgt aaa     384
Val Glu Thr Gly Met Ala Trp Phe Trp Pro Asp Ser Arg Val Arg Lys
        115                 120                 125

tgg ttt gat gtt gaa ggg ttg gat aac ctt aaa cgc gca caa atg caa     432
Trp Phe Asp Val Glu Gly Leu Asp Asn Leu Lys Arg Ala Gln Met Gln
    130                 135                 140

aat cgc ggc gta atg gtt gtc ggc gtc cat ttt atg tcg ctg gaa ctg     480
Asn Arg Gly Val Met Val Val Gly Val His Phe Met Ser Leu Glu Leu
145                 150                 155                 160

ggc ggc cgc gtg atg gga ctg tgc caa cca atg atg gct acc tat cgt     528
Gly Gly Arg Val Met Gly Leu Cys Gln Pro Met Met Ala Thr Tyr Arg
                165                 170                 175

cca cat aat aat cag ctg atg gaa tgg gtg cag acc cgt ggg cgc atg     576
Pro His Asn Asn Gln Leu Met Glu Trp Val Gln Thr Arg Gly Arg Met
            180                 185                 190

cgc tct aac aaa gcg atg atc ggc aga aat aat ctg cgc ggc att gtc     624
Arg Ser Asn Lys Ala Met Ile Gly Arg Asn Asn Leu Arg Gly Ile Val
        195                 200                 205

ggt gca ctg aag aaa ggt gaa gcg gta tgg ttt gct ccc gat cag gat     672
Gly Ala Leu Lys Lys Gly Glu Ala Val Trp Phe Ala Pro Asp Gln Asp
    210                 215                 220

tat ggt cgt aaa ggc agc tcc ttc gcg ccg ttc ttt gcg gtg gaa aat     720
Tyr Gly Arg Lys Gly Ser Ser Phe Ala Pro Phe Phe Ala Val Glu Asn
225                 230                 235                 240

gtc gcc aca acc aat ggc acc tat gtt ctc tcc cgt ctc tct ggc gca     768
Val Ala Thr Thr Asn Gly Thr Tyr Val Leu Ser Arg Leu Ser Gly Ala
                245                 250                 255

gcc atg ttg acc gta acg atg gta aga aaa gcg gat tac agc gga tat     816
Ala Met Leu Thr Val Thr Met Val Arg Lys Ala Asp Tyr Ser Gly Tyr
            260                 265                 270

cgt ttg ttc atc acc cca gag atg gaa ggc tac ccg aca gat gaa aat     864
Arg Leu Phe Ile Thr Pro Glu Met Glu Gly Tyr Pro Thr Asp Glu Asn
        275                 280                 285

caa gcc gct gcc tat atg aac aag att atc gag aaa gag atc atg cgc     912
Gln Ala Ala Ala Tyr Met Asn Lys Ile Ile Glu Lys Glu Ile Met Arg
    290                 295                 300

gca ccg gag cag tac ctc tgg atc cac cgt cgc ttt aaa acg cgc ccg     960
Ala Pro Glu Gln Tyr Leu Trp Ile His Arg Arg Phe Lys Thr Arg Pro
305                 310                 315                 320

gtg gga gaa tcg tcg ttg tac att taa                                 987
Val Gly Glu Ser Ser Leu Tyr Ile  *
                325
```

<210> 248
<211> 264
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(264)

<400> 248
ttg gct aat atc aaa tca gct aag aag cgc gcc att cag tct gaa aag 48
Met Ala Asn Ile Lys Ser Ala Lys Lys Arg Ala Ile Gln Ser Glu Lys
1 5 10 15

gct cgt aag cac aac gca agc cgt cgc tct atg atg cgt act ttc atc 96
Ala Arg Lys His Asn Ala Ser Arg Arg Ser Met Met Arg Thr Phe Ile
20 25 30

aag aaa gta tac gca gct atc gaa gct ggc gac aaa gct gct gca cag 144
Lys Lys Val Tyr Ala Ala Ile Glu Ala Gly Asp Lys Ala Ala Ala Gln
35 40 45

aaa gca ttt aac gaa atg caa ccg atc gtg gac cgt cag gct gct aaa 192
Lys Ala Phe Asn Glu Met Gln Pro Ile Val Asp Arg Gln Ala Ala Lys
50 55 60

ggt ctg atc cac aaa aac aaa gct gca cgt cat aag gct aac ctg act 240
Gly Leu Ile His Lys Asn Lys Ala Ala Arg His Lys Ala Asn Leu Thr
65 70 75 80

gca cag atc aac aaa ctg gct taa 264
Ala Gln Ile Asn Lys Leu Ala *
85

<210> 249
<211> 1293
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1293)

<400> 249
atg cgc tat aat ggt tta aat aat atg ttt ttc cct ctt tgc ctg att 48
Met Arg Tyr Asn Gly Leu Asn Asn Met Phe Phe Pro Leu Cys Leu Ile
1 5 10 15

aac gat aac cac tct gtc aca agt cca tca cat aca aag aaa aca aaa 96
Asn Asp Asn His Ser Val Thr Ser Pro Ser His Thr Lys Lys Thr Lys
20 25 30

tca gat aat tac agc aaa cat cat aaa aac acg tta att gac aat aaa 144
Ser Asp Asn Tyr Ser Lys His His Lys Asn Thr Leu Ile Asp Asn Lys
35 40 45

gcc ctc tct ctt ttc aaa atg gat gat cat gaa aaa gtg ata ggc ttg 192
Ala Leu Ser Leu Phe Lys Met Asp Asp His Glu Lys Val Ile Gly Leu
50 55 60

```
att cag aaa atg aaa aga att tat gat agt tta cca tca gga aaa atc    240
Ile Gln Lys Met Lys Arg Ile Tyr Asp Ser Leu Pro Ser Gly Lys Ile
 65                 70                  75                  80

acg aaa gaa acg gac agg aaa ata cat aaa tat ttt ata gat ata gct    288
Thr Lys Glu Thr Asp Arg Lys Ile His Lys Tyr Phe Ile Asp Ile Ala
                85                  90                  95

tca cat gca aat aat aaa tgt gac gat aga att acg aga aga gtt tac    336
Ser His Ala Asn Asn Lys Cys Asp Asp Arg Ile Thr Arg Arg Val Tyr
            100                 105                 110

ctt aat aaa gat aag gaa gtg tca att aag gtg gta tat ttt ata aat    384
Leu Asn Lys Asp Lys Glu Val Ser Ile Lys Val Val Tyr Phe Ile Asn
        115                 120                 125

aat gtc acc gtc cat aat aat act atc gaa atc cca cag aca gta aat    432
Asn Val Thr Val His Asn Asn Thr Ile Glu Ile Pro Gln Thr Val Asn
    130                 135                 140

ggt ggt tac gat ttt tca cac ctt agc ctg aaa ggt atc gtg att aaa    480
Gly Gly Tyr Asp Phe Ser His Leu Ser Leu Lys Gly Ile Val Ile Lys
145                 150                 155                 160

gat gaa gat tta tcc aat tcg aat ttt gca ggt tgc aga cta caa aac    528
Asp Glu Asp Leu Ser Asn Ser Asn Phe Ala Gly Cys Arg Leu Gln Asn
                165                 170                 175

gct att ttt cag gac tgt aat atg tat aaa acg aat ttt aat ttc gcc    576
Ala Ile Phe Gln Asp Cys Asn Met Tyr Lys Thr Asn Phe Asn Phe Ala
            180                 185                 190

ata atg gaa aaa ata ctt ttt gat aat tgt att ctc gat gac tca aat    624
Ile Met Glu Lys Ile Leu Phe Asp Asn Cys Ile Leu Asp Asp Ser Asn
        195                 200                 205

ttc gct cag ata aaa atg act gac gga act cta aat tca tgt tcc gct    672
Phe Ala Gln Ile Lys Met Thr Asp Gly Thr Leu Asn Ser Cys Ser Ala
    210                 215                 220

atg cat gtt caa ttc tac aat gca aca atg aat aga gcc aat att aaa    720
Met His Val Gln Phe Tyr Asn Ala Thr Met Asn Arg Ala Asn Ile Lys
225                 230                 235                 240

aat acc ttc ctt gat tat tca aat ttt tat atg gca tac atg gct gag    768
Asn Thr Phe Leu Asp Tyr Ser Asn Phe Tyr Met Ala Tyr Met Ala Glu
                245                 250                 255

gta aat ctt tat aaa gta ata gcg cca tat att aat tta ttt aga gcc    816
Val Asn Leu Tyr Lys Val Ile Ala Pro Tyr Ile Asn Leu Phe Arg Ala
            260                 265                 270

gac ctt agc ttc tct aaa ctt gat tta att aac ttt gaa cat gct gat    864
Asp Leu Ser Phe Ser Lys Leu Asp Leu Ile Asn Phe Glu His Ala Asp
        275                 280                 285

ctg tct cgt gtc aac ctg aat aaa gca acc ctc cag aat ata aac tta    912
Leu Ser Arg Val Asn Leu Asn Lys Ala Thr Leu Gln Asn Ile Asn Leu
    290                 295                 300
```

att gat agc aaa ctc ttt ttt acg cgg tta aca aat acg ttc ctc gaa 960
Ile Asp Ser Lys Leu Phe Phe Thr Arg Leu Thr Asn Thr Phe Leu Glu
305 310 315 320

atg gtt ata tgt acc gac tct aat atg gct aat gtt aat ttt aat aat 1008
Met Val Ile Cys Thr Asp Ser Asn Met Ala Asn Val Asn Phe Asn Asn
325 330 335

gcc aat tta agc aat tgc cat ttc aac tgt tct gtt tta aca aaa gcc 1056
Ala Asn Leu Ser Asn Cys His Phe Asn Cys Ser Val Leu Thr Lys Ala
340 345 350

tgg atg ttt aat atc cgt ctc tat cgt gtt aat ttc gat gag gct agc 1104
Trp Met Phe Asn Ile Arg Leu Tyr Arg Val Asn Phe Asp Glu Ala Ser
355 360 365

gtc cag gga atg ggt att acc att ctc cgt ggt gag gaa aat atc tcc 1152
Val Gln Gly Met Gly Ile Thr Ile Leu Arg Gly Glu Glu Asn Ile Ser
370 375 380

att aat agt gat atc ctg gta aca cta cag aaa ttc ttt gaa gaa gat 1200
Ile Asn Ser Asp Ile Leu Val Thr Leu Gln Lys Phe Phe Glu Glu Asp
385 390 395 400

tgt gcc act cat act ggc atg tca caa act gag gat aat ctt cat gca 1248
Cys Ala Thr His Thr Gly Met Ser Gln Thr Glu Asp Asn Leu His Ala
405 410 415

gtc gct atg aag att act gca gat att atg caa gat gca gat tga 1293
Val Ala Met Lys Ile Thr Ala Asp Ile Met Gln Asp Ala Asp *
420 425 430

<210> 250
<211> 2652
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(2652)

<400> 250
atg aac gaa caa tat tcc gca ttg cgt agt aat gtc agt atg ctc ggc 48
Met Asn Glu Gln Tyr Ser Ala Leu Arg Ser Asn Val Ser Met Leu Gly
1 5 10 15

aaa gtg ctg gga gaa acc atc aag gat gcg ttg gga gaa cac att ctt 96
Lys Val Leu Gly Glu Thr Ile Lys Asp Ala Leu Gly Glu His Ile Leu
20 25 30

gaa cgc gta gaa act atc cgt aag ttg tcg aaa tct tca cgc gct ggc 144
Glu Arg Val Glu Thr Ile Arg Lys Leu Ser Lys Ser Ser Arg Ala Gly
35 40 45

aat gat gct aac cgc cag gag ttg ctc acc acc tta caa aat ttg tcg 192
Asn Asp Ala Asn Arg Gln Glu Leu Leu Thr Thr Leu Gln Asn Leu Ser
50 55 60

aac gac gag ctg ctg ccc gtt gcg cgt gcg ttt agt cag ttc ctg aac 240

```
Asn Asp Glu Leu Leu Pro Val Ala Arg Ala Phe Ser Gln Phe Leu Asn
 65                  70                  75                  80

ctg gcc aac acc gcc gag caa tac cac agc att tcg ccg aaa ggc gaa     288
Leu Ala Asn Thr Ala Glu Gln Tyr His Ser Ile Ser Pro Lys Gly Glu
                85                  90                  95

gct gcc agc aac ccg gaa gtg atc gcc cgc acc ctg cgt aaa ctg aaa     336
Ala Ala Ser Asn Pro Glu Val Ile Ala Arg Thr Leu Arg Lys Leu Lys
            100                 105                 110

aac cag ccg gaa ctg agc gaa gac acc atc aaa aaa gca gtg gaa tcg     384
Asn Gln Pro Glu Leu Ser Glu Asp Thr Ile Lys Lys Ala Val Glu Ser
        115                 120                 125

ctg tcg ctg gaa ctg gtc ctc acg gct cac cca acc gaa att acc cgt     432
Leu Ser Leu Glu Leu Val Leu Thr Ala His Pro Thr Glu Ile Thr Arg
    130                 135                 140

cgt aca ctg atc cac aaa atg gtg gaa gtg aac gcc tgt tta aaa cag     480
Arg Thr Leu Ile His Lys Met Val Glu Val Asn Ala Cys Leu Lys Gln
145                 150                 155                 160

ctc gat aac aaa gat atc gct gac tac gaa cac aac cag ctg atg cgt     528
Leu Asp Asn Lys Asp Ile Ala Asp Tyr Glu His Asn Gln Leu Met Arg
                165                 170                 175

cgc ctg cgc cag ttg atc gcc cag tca tgg cat acc gat gaa atc cgt     576
Arg Leu Arg Gln Leu Ile Ala Gln Ser Trp His Thr Asp Glu Ile Arg
            180                 185                 190

aag ctg cgt cca agc ccg gta gat gaa gcc aaa tgg ggc ttt gcc gta     624
Lys Leu Arg Pro Ser Pro Val Asp Glu Ala Lys Trp Gly Phe Ala Val
        195                 200                 205

gtg gaa aac agc ctg tgg caa ggc gta cca aat tac ctg cgc gaa ctg     672
Val Glu Asn Ser Leu Trp Gln Gly Val Pro Asn Tyr Leu Arg Glu Leu
    210                 215                 220

aac gaa caa ctg gaa gag aac ctc ggc tac aaa ctg ccc gtc gaa ttt     720
Asn Glu Gln Leu Glu Glu Asn Leu Gly Tyr Lys Leu Pro Val Glu Phe
225                 230                 235                 240

gtt ccg gtc cgt ttt act tcg tgg atg ggc ggc gac cgc gac ggc aac     768
Val Pro Val Arg Phe Thr Ser Trp Met Gly Gly Asp Arg Asp Gly Asn
                245                 250                 255

ccg aac gtc act gcc gat atc acc cgc cac gtc ctg cta ctc agc cgc     816
Pro Asn Val Thr Ala Asp Ile Thr Arg His Val Leu Leu Leu Ser Arg
            260                 265                 270

tgg aaa gcc acc gat ttg ttc ctg aaa gat att cag gtg ctg gtt tct     864
Trp Lys Ala Thr Asp Leu Phe Leu Lys Asp Ile Gln Val Leu Val Ser
        275                 280                 285

gaa ctg tcg atg gtt gaa gcg acc cct gaa ctg ctg gcg ctg gtt ggc     912
Glu Leu Ser Met Val Glu Ala Thr Pro Glu Leu Leu Ala Leu Val Gly
    290                 295                 300

gaa gaa ggt gcc gca gaa ccg tat cgc tat ctg atg aaa aac ctg cgt     960
Glu Glu Gly Ala Ala Glu Pro Tyr Arg Tyr Leu Met Lys Asn Leu Arg
```

305 310 315 320

tct cgc ctg atg gcg aca cag gca tgg ctg gaa gcg cgc ctg aaa ggc 1008
Ser Arg Leu Met Ala Thr Gln Ala Trp Leu Glu Ala Arg Leu Lys Gly
325 330 335

gaa gaa ctg cca aaa cca gaa ggc ctg ctg aca caa aac gaa gaa ctg 1056
Glu Glu Leu Pro Lys Pro Glu Gly Leu Leu Thr Gln Asn Glu Glu Leu
340 345 350

tgg gaa ccg ctc tac gct tgc tac cag tca ctt cag gcg tgt ggc atg 1104
Trp Glu Pro Leu Tyr Ala Cys Tyr Gln Ser Leu Gln Ala Cys Gly Met
355 360 365

ggt att atc gcc aac ggc gat ctg ctc gac acc ctg cgc cgc gtg aaa 1152
Gly Ile Ile Ala Asn Gly Asp Leu Leu Asp Thr Leu Arg Arg Val Lys
370 375 380

tgt ttc ggc gta ccg ctg gtc cgt att gat atc cgt cag gag agc acg 1200
Cys Phe Gly Val Pro Leu Val Arg Ile Asp Ile Arg Gln Glu Ser Thr
385 390 395 400

cgt cat acc gaa gcg ctg ggc gag ctg acc cgc tac ctc ggt atc ggc 1248
Arg His Thr Glu Ala Leu Gly Glu Leu Thr Arg Tyr Leu Gly Ile Gly
405 410 415

gac tac gaa agc tgg tca gag gcc gac aaa cag gcg ttc ctg atc cgc 1296
Asp Tyr Glu Ser Trp Ser Glu Ala Asp Lys Gln Ala Phe Leu Ile Arg
420 425 430

gaa ctg aac tcc aaa cgt ccg ctt ctg ccg cgc aac tgg caa cca agc 1344
Glu Leu Asn Ser Lys Arg Pro Leu Leu Pro Arg Asn Trp Gln Pro Ser
435 440 445

gcc gaa acg cgc gaa gtg ctc gat acc tgc cag gtg att gcc gaa gca 1392
Ala Glu Thr Arg Glu Val Leu Asp Thr Cys Gln Val Ile Ala Glu Ala
450 455 460

ccg caa ggc tcc att gcc gcc tac gtg atc tcg atg gcg aaa acg ccg 1440
Pro Gln Gly Ser Ile Ala Ala Tyr Val Ile Ser Met Ala Lys Thr Pro
465 470 475 480

tcc gac gta ctg gct gtc cac ctg ctg ctg aaa gaa gcg ggt atc ggg 1488
Ser Asp Val Leu Ala Val His Leu Leu Leu Lys Glu Ala Gly Ile Gly
485 490 495

ttt gcg atg ccg gtt gct ccg ctg ttt gaa acc ctc gat gat ctg aac 1536
Phe Ala Met Pro Val Ala Pro Leu Phe Glu Thr Leu Asp Asp Leu Asn
500 505 510

aac gcc aac gat gtc atg acc cag ctg ctc aat att gac tgg tat cgt 1584
Asn Ala Asn Asp Val Met Thr Gln Leu Leu Asn Ile Asp Trp Tyr Arg
515 520 525

ggc ctg att cag ggc aaa cag atg gtg atg att ggc tat tcc gac tca 1632
Gly Leu Ile Gln Gly Lys Gln Met Val Met Ile Gly Tyr Ser Asp Ser
530 535 540

gca aaa gat gcg gga gtg atg gca gct tcc tgg gcg caa tat cag gca 1680
Ala Lys Asp Ala Gly Val Met Ala Ala Ser Trp Ala Gln Tyr Gln Ala
545 550 555 560

cag gat gca tta atc aaa acc tgc gaa aaa gcg ggt att gag ctg acg 1728
Gln Asp Ala Leu Ile Lys Thr Cys Glu Lys Ala Gly Ile Glu Leu Thr
565 570 575

ttg ttc cac ggt cgc ggc ggt tcc att ggt cgc ggc ggc gca cct gct 1776
Leu Phe His Gly Arg Gly Gly Ser Ile Gly Arg Gly Gly Ala Pro Ala
580 585 590

cat gcg gcg ctg ctg tca caa ccg cca gga agc ctg aaa ggc ggc ctg 1824
His Ala Ala Leu Leu Ser Gln Pro Pro Gly Ser Leu Lys Gly Gly Leu
595 600 605

cgc gta acc gaa cag ggc gag atg atc cgc ttt aaa tat ggt ctg cca 1872
Arg Val Thr Glu Gln Gly Glu Met Ile Arg Phe Lys Tyr Gly Leu Pro
610 615 620

gaa atc acc gtc agc agc ctg tcg ctt tat acc ggg gcg att ctg gaa 1920
Glu Ile Thr Val Ser Ser Leu Ser Leu Tyr Thr Gly Ala Ile Leu Glu
625 630 635 640

gcc aac ctg ctg cca ccg ccg gag ccg aaa gag agc tgg cgt cgc att 1968
Ala Asn Leu Leu Pro Pro Pro Glu Pro Lys Glu Ser Trp Arg Arg Ile
645 650 655

atg gat gaa ctg tca gtc atc tcc tgc gat gtc tac cgc ggc tac gta 2016
Met Asp Glu Leu Ser Val Ile Ser Cys Asp Val Tyr Arg Gly Tyr Val
660 665 670

cgt gaa aac aaa gat ttt gtg cct tac ttc cgc tcc gct acg ccg gaa 2064
Arg Glu Asn Lys Asp Phe Val Pro Tyr Phe Arg Ser Ala Thr Pro Glu
675 680 685

caa gaa ctg ggc aaa ctg ccg ttg ggt tca cgt ccg gcg aaa cgt cgc 2112
Gln Glu Leu Gly Lys Leu Pro Leu Gly Ser Arg Pro Ala Lys Arg Arg
690 695 700

cca acc ggc ggc gtc gag tca cta cgc gcc att ccg tgg atc ttc gcc 2160
Pro Thr Gly Gly Val Glu Ser Leu Arg Ala Ile Pro Trp Ile Phe Ala
705 710 715 720

tgg acg caa aac cgt ctg atg ctc ccc gcc tgg ctg ggt gca ggt acg 2208
Trp Thr Gln Asn Arg Leu Met Leu Pro Ala Trp Leu Gly Ala Gly Thr
725 730 735

gcg ctg caa aaa gtg gtc gaa gac ggc aaa cag agc gag ctg gag gct 2256
Ala Leu Gln Lys Val Val Glu Asp Gly Lys Gln Ser Glu Leu Glu Ala
740 745 750

atg tgc cgc gat tgg cca ttc ttc tcg acg cgt ctc ggc atg ctg gag 2304
Met Cys Arg Asp Trp Pro Phe Phe Ser Thr Arg Leu Gly Met Leu Glu
755 760 765

atg gtc ttc gcc aaa gca gac ctg tgg ctg gcg gaa tac tat gac caa 2352
Met Val Phe Ala Lys Ala Asp Leu Trp Leu Ala Glu Tyr Tyr Asp Gln
770 775 780

cgc ctg gta gac aaa gca ctg tgg ccg tta ggt aaa gag tta cgc aac 2400
Arg Leu Val Asp Lys Ala Leu Trp Pro Leu Gly Lys Glu Leu Arg Asn
785 790 795 800

ctg caa gaa gaa gac atc aaa gtg gtg ctg gcg att gcc aac gat tcc 2448
Leu Gln Glu Glu Asp Ile Lys Val Val Leu Ala Ile Ala Asn Asp Ser
805 810 815

cat ctg atg gcc gat ctg ccg tgg att gca gag tct att cag cta cgg 2496
His Leu Met Ala Asp Leu Pro Trp Ile Ala Glu Ser Ile Gln Leu Arg
820 825 830

aat att tac acc gac ccg ctg aac gta ttg cag gcc gag ttg ctg cac 2544
Asn Ile Tyr Thr Asp Pro Leu Asn Val Leu Gln Ala Glu Leu Leu His
835 840 845

cgc tcc cgc cag gca gaa aaa gaa ggc cag gaa ccg gat cct cgc gtc 2592
Arg Ser Arg Gln Ala Glu Lys Glu Gly Gln Glu Pro Asp Pro Arg Val
850 855 860

gaa caa gcg tta atg gtc act att gcc ggg att gcg gca ggt atg cgt 2640
Glu Gln Ala Leu Met Val Thr Ile Ala Gly Ile Ala Ala Gly Met Arg
865 870 875 880

aat acc ggc taa 2652
Asn Thr Gly *

<210> 251
<211> 1992
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1992)

<400> 251
atg tcc tca cgt aaa gag ctt gcc aat gct att cgt gcg ctg agc atg 48
Met Ser Ser Arg Lys Glu Leu Ala Asn Ala Ile Arg Ala Leu Ser Met
1 5 10 15

gac gca gta cag aaa gcc aaa tcc ggt cac ccg ggt gcc cct atg ggt 96
Asp Ala Val Gln Lys Ala Lys Ser Gly His Pro Gly Ala Pro Met Gly
20 25 30

atg gct gac att gcc gaa gtc ctg tgg cgt gat ttc ctg aaa cac aac 144
Met Ala Asp Ile Ala Glu Val Leu Trp Arg Asp Phe Leu Lys His Asn
35 40 45

ccg cag aat ccg tcc tgg gct gac cgt gac cgc ttc gtg ctg tcc aac 192
Pro Gln Asn Pro Ser Trp Ala Asp Arg Asp Arg Phe Val Leu Ser Asn
50 55 60

ggc cac ggc tcc atg ctg atc tac agc ctg ctg cac ctc acc ggt tac 240
Gly His Gly Ser Met Leu Ile Tyr Ser Leu Leu His Leu Thr Gly Tyr
65 70 75 80

gat ctg ccg atg gaa gaa ctg aaa aac ttc cgt cag ctg cac tct aaa 288
Asp Leu Pro Met Glu Glu Leu Lys Asn Phe Arg Gln Leu His Ser Lys
85 90 95

act ccg ggt cac ccg gaa gtg ggt tac acc gct ggt gtg gaa acc acc 336
Thr Pro Gly His Pro Glu Val Gly Tyr Thr Ala Gly Val Glu Thr Thr

100 105 110

acc ggt ccg ctg ggt cag ggt att gcc aac gca gtc ggt atg gcg att 384
Thr Gly Pro Leu Gly Gln Gly Ile Ala Asn Ala Val Gly Met Ala Ile
115 120 125

gca gaa aaa acg ctg gcg gcg cag ttt aac cgt ccg ggc cac gac att 432
Ala Glu Lys Thr Leu Ala Ala Gln Phe Asn Arg Pro Gly His Asp Ile
130 135 140

gtc gac cac tac acc tac gcc ttc atg ggc gac ggc tgc atg atg gaa 480
Val Asp His Tyr Thr Tyr Ala Phe Met Gly Asp Gly Cys Met Met Glu
145 150 155 160

ggc atc tcc cac gaa gtt tgc tct ctg gcg ggt acg ctg aag ctg ggt 528
Gly Ile Ser His Glu Val Cys Ser Leu Ala Gly Thr Leu Lys Leu Gly
165 170 175

aaa ctg att gca ttc tac gat gac aac ggt att tct atc gat ggt cac 576
Lys Leu Ile Ala Phe Tyr Asp Asp Asn Gly Ile Ser Ile Asp Gly His
180 185 190

gtt gaa ggc tgg ttc acc gac gac acc gca atg cgt ttc gaa gct tac 624
Val Glu Gly Trp Phe Thr Asp Asp Thr Ala Met Arg Phe Glu Ala Tyr
195 200 205

ggc tgg cac gtt att cgc gac atc gac ggt cat gac gcg gca tct atc 672
Gly Trp His Val Ile Arg Asp Ile Asp Gly His Asp Ala Ala Ser Ile
210 215 220

aaa cgc gca gta gaa gaa gcg cgc gca gtg act gac aaa cct tcc ctg 720
Lys Arg Ala Val Glu Glu Ala Arg Ala Val Thr Asp Lys Pro Ser Leu
225 230 235 240

ctg atg tgc aaa acc atc atc ggt ttc ggt tcc ccg aac aaa gcc ggt 768
Leu Met Cys Lys Thr Ile Ile Gly Phe Gly Ser Pro Asn Lys Ala Gly
245 250 255

acc cac gac tcc cac ggt gcg ccg ctg ggc gac gct gaa att gcc ctg 816
Thr His Asp Ser His Gly Ala Pro Leu Gly Asp Ala Glu Ile Ala Leu
260 265 270

acc cgc gaa caa ctg ggc tgg aaa tat gcg ccg ttc gaa atc ccg tct 864
Thr Arg Glu Gln Leu Gly Trp Lys Tyr Ala Pro Phe Glu Ile Pro Ser
275 280 285

gaa atc tat gct cag tgg gat gcg aaa gaa gca ggc cag gcg aaa gaa 912
Glu Ile Tyr Ala Gln Trp Asp Ala Lys Glu Ala Gly Gln Ala Lys Glu
290 295 300

tcc gca tgg aac gag aaa ttc gct gct tac gcg aaa gct tat ccg cag 960
Ser Ala Trp Asn Glu Lys Phe Ala Ala Tyr Ala Lys Ala Tyr Pro Gln
305 310 315 320

gaa gcc gct gaa ttt acc cgc cgt atg aaa ggc gaa atg ccg tct gac 1008
Glu Ala Ala Glu Phe Thr Arg Arg Met Lys Gly Glu Met Pro Ser Asp
325 330 335

ttc gac gct aaa gcg aaa gag ttc atc gct aaa ctg cag gct aat ccg 1056
Phe Asp Ala Lys Ala Lys Glu Phe Ile Ala Lys Leu Gln Ala Asn Pro
340 345 350

```
gcg aaa atc gcc agc cgt aaa gcg tct cag aat gct atc gaa gcg ttc      1104
Ala Lys Ile Ala Ser Arg Lys Ala Ser Gln Asn Ala Ile Glu Ala Phe
        355                 360                 365

ggt ccg ctg ttg ccg gaa ttc ctc ggc ggt tct gct gac ctg gcg ccg      1152
Gly Pro Leu Leu Pro Glu Phe Leu Gly Gly Ser Ala Asp Leu Ala Pro
    370                 375                 380

tct aac ctg acc ctg tgg tct ggt tct aaa gca atc aac gaa gat gct      1200
Ser Asn Leu Thr Leu Trp Ser Gly Ser Lys Ala Ile Asn Glu Asp Ala
385                 390                 395                 400

gcg ggt aac tac atc cac tac ggt gtt cgc gag ttc ggt atg acc gcg      1248
Ala Gly Asn Tyr Ile His Tyr Gly Val Arg Glu Phe Gly Met Thr Ala
                405                 410                 415

att gct aac ggt atc tcc ctg cac ggt ggc ttc ctg ccg tac acc tcc      1296
Ile Ala Asn Gly Ile Ser Leu His Gly Gly Phe Leu Pro Tyr Thr Ser
            420                 425                 430

acc ttc ctg atg ttc gtg gaa tac gca cgt aac gcc gta cgt atg gct      1344
Thr Phe Leu Met Phe Val Glu Tyr Ala Arg Asn Ala Val Arg Met Ala
        435                 440                 445

gcg ctg atg aaa cag cgt cag gtg atg gtt tac acc cac gac tcc atc      1392
Ala Leu Met Lys Gln Arg Gln Val Met Val Tyr Thr His Asp Ser Ile
    450                 455                 460

ggt ctg ggc gaa gac ggg ccg act cac cag ccg gtt gag cag gtc gct      1440
Gly Leu Gly Glu Asp Gly Pro Thr His Gln Pro Val Glu Gln Val Ala
465                 470                 475                 480

tct ctg cgc gta acc ccg aac atg tct aca tgg cgt ccg tgt gac cag      1488
Ser Leu Arg Val Thr Pro Asn Met Ser Thr Trp Arg Pro Cys Asp Gln
                485                 490                 495

gtt gaa tcc gcg gtc gcg tgg aaa tac ggt gtt gag cgt cag gac ggc      1536
Val Glu Ser Ala Val Ala Trp Lys Tyr Gly Val Glu Arg Gln Asp Gly
            500                 505                 510

ccg acc gca ctg atc ctc tcc cgt cag aac ctg gcg cag cag gaa cga      1584
Pro Thr Ala Leu Ile Leu Ser Arg Gln Asn Leu Ala Gln Gln Glu Arg
        515                 520                 525

act gaa gag caa ctg gca aac atc gcg cgc ggt ggt tat gtg ctg aaa      1632
Thr Glu Glu Gln Leu Ala Asn Ile Ala Arg Gly Gly Tyr Val Leu Lys
    530                 535                 540

gac tgc gcc ggt cag ccg gaa ctg att ttc atc gct acc ggt tca gaa      1680
Asp Cys Ala Gly Gln Pro Glu Leu Ile Phe Ile Ala Thr Gly Ser Glu
545                 550                 555                 560

gtt gaa ctg gct gtt gct gcc tac gaa aaa ctg act gcc gaa ggc gtg      1728
Val Glu Leu Ala Val Ala Ala Tyr Glu Lys Leu Thr Ala Glu Gly Val
                565                 570                 575

aaa gcg cgc gtg gtg tcc atg tcg tct acc gac gca ttt gac aag cag      1776
Lys Ala Arg Val Val Ser Met Ser Ser Thr Asp Ala Phe Asp Lys Gln
            580                 585                 590
```

```
gat gct gct tac cgt gaa tcc gta ctg ccg aaa gcg gtt act gca cgc    1824
Asp Ala Ala Tyr Arg Glu Ser Val Leu Pro Lys Ala Val Thr Ala Arg
        595                 600                 605

gtt gct gta gaa gcg ggt att gct gac tac tgg tac aag tat gtt ggc    1872
Val Ala Val Glu Ala Gly Ile Ala Asp Tyr Trp Tyr Lys Tyr Val Gly
    610                 615                 620

ctg aac ggt gct atc gtc ggt atg acc acc ttc ggt gaa tct gct ccg    1920
Leu Asn Gly Ala Ile Val Gly Met Thr Thr Phe Gly Glu Ser Ala Pro
625                 630                 635                 640

gca gag ctg ctg ttt gaa gag ttc ggc ttc act gtt gat aac gtt gtt    1968
Ala Glu Leu Leu Phe Glu Glu Phe Gly Phe Thr Val Asp Asn Val Val
                645                 650                 655

gcg aaa gca aaa gaa ctg ctg taa                                    1992
Ala Lys Ala Lys Glu Leu Leu  *
            660


<210> 252
<211> 459
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(459)

<400> 252
atg ttc cgg gga gca acg tta gtc aat ctc gac agc aaa ggg cgc tta    48
Met Phe Arg Gly Ala Thr Leu Val Asn Leu Asp Ser Lys Gly Arg Leu
 1               5                  10                  15

tca gtg cct acc cgt tat cgg gaa cag ctg ctt gag aac gct gcc ggt    96
Ser Val Pro Thr Arg Tyr Arg Glu Gln Leu Leu Glu Asn Ala Ala Gly
            20                  25                  30

caa atg gtt tgc acc att gac att tat cac ccg tgc ctg ctg ctt tac    144
Gln Met Val Cys Thr Ile Asp Ile Tyr His Pro Cys Leu Leu Leu Tyr
        35                  40                  45

ccc ctg cct gaa tgg gaa att atc gag caa aaa tta tcg cgt ctg tcg    192
Pro Leu Pro Glu Trp Glu Ile Ile Glu Gln Lys Leu Ser Arg Leu Ser
    50                  55                  60

agc atg aac ccg gtt gag cgc cgt gtg cag cgc cta ctg tta ggt cat    240
Ser Met Asn Pro Val Glu Arg Arg Val Gln Arg Leu Leu Leu Gly His
65                  70                  75                  80

gcc agc gaa tgt cag atg gat ggc gca ggt cga ttg tta atc gcg cca    288
Ala Ser Glu Cys Gln Met Asp Gly Ala Gly Arg Leu Leu Ile Ala Pro
                85                  90                  95

gta ctg cgg caa cat gcc ggg ctg aca aaa gaa gtg atg ctg gtt gga    336
Val Leu Arg Gln His Ala Gly Leu Thr Lys Glu Val Met Leu Val Gly
            100                 105                 110

cag ttc aac aag ttt gag ctg tgg gat gaa aca acc tgg cat caa cag    384
Gln Phe Asn Lys Phe Glu Leu Trp Asp Glu Thr Thr Trp His Gln Gln
```

115 120 125

gtc aag gaa gat atc gac gca gag cag ttg gct acc gga gac tta tcg 432
Val Lys Glu Asp Ile Asp Ala Glu Gln Leu Ala Thr Gly Asp Leu Ser
130 135 140

gag cga ctg cag gac ttg tct cta taa 459
Glu Arg Leu Gln Asp Leu Ser Leu *
145 150

<210> 253
<211> 942
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(942)

<400> 253
atg atg gaa aac tat aaa cat act acg gtg ctg ctg gat gaa gcc gtt 48
Met Met Glu Asn Tyr Lys His Thr Thr Val Leu Leu Asp Glu Ala Val
1 5 10 15

aat ggc ctc aat atc cgt cct gat ggc atc tac att gat ggg act ttt 96
Asn Gly Leu Asn Ile Arg Pro Asp Gly Ile Tyr Ile Asp Gly Thr Phe
20 25 30

ggt cgc ggt ggt cac tca cgt ctg atc ctc tcg cag ctt ggc gaa gag 144
Gly Arg Gly Gly His Ser Arg Leu Ile Leu Ser Gln Leu Gly Glu Glu
35 40 45

ggg cgt ttg ctg gcg atc gat cgc gac ccg cag gct atc gcc gtt gcg 192
Gly Arg Leu Leu Ala Ile Asp Arg Asp Pro Gln Ala Ile Ala Val Ala
50 55 60

aag act att gat gat ccg cgc ttc tcc atc atc cac gga cct ttc tcc 240
Lys Thr Ile Asp Asp Pro Arg Phe Ser Ile Ile His Gly Pro Phe Ser
65 70 75 80

gcg ctg ggc gaa tac gtt gcc gag cgc gat ctt atc ggc aag atc gac 288
Ala Leu Gly Glu Tyr Val Ala Glu Arg Asp Leu Ile Gly Lys Ile Asp
85 90 95

ggc att ctc ctc gat ctt ggc gtc tct tca ccg caa ctt gat gat gct 336
Gly Ile Leu Leu Asp Leu Gly Val Ser Ser Pro Gln Leu Asp Asp Ala
100 105 110

gaa cgt ggc ttt tcc ttt atg cgc gat ggt ccg ctg gac atg cgt atg 384
Glu Arg Gly Phe Ser Phe Met Arg Asp Gly Pro Leu Asp Met Arg Met
115 120 125

gac cca acc cgt ggg cag tca gcc gct gaa tgg cta caa acc gca gaa 432
Asp Pro Thr Arg Gly Gln Ser Ala Ala Glu Trp Leu Gln Thr Ala Glu
130 135 140

gaa gcc gat atc gcc tgg gta ttg aaa acc tat ggt gaa gag cgt ttt 480
Glu Ala Asp Ile Ala Trp Val Leu Lys Thr Tyr Gly Glu Glu Arg Phe
145 150 155 160

```
gcc aaa cgc att gcc cgc gcc att gtc gag cgt aac cgc gaa cag ccg     528
Ala Lys Arg Ile Ala Arg Ala Ile Val Glu Arg Asn Arg Glu Gln Pro
                165                 170                 175

atg acc cgc acc aaa gaa ctg gcg gaa gtc gtg gct gct gca acg ccg     576
Met Thr Arg Thr Lys Glu Leu Ala Glu Val Val Ala Ala Ala Thr Pro
            180                 185                 190

gtg aaa gat aag ttt aaa cat ccc gcg acc cgt acc ttc cag gcg gtg     624
Val Lys Asp Lys Phe Lys His Pro Ala Thr Arg Thr Phe Gln Ala Val
        195                 200                 205

cgc att tgg gta aac agt gaa ctg gag gag ata gag cag gcg cta aaa     672
Arg Ile Trp Val Asn Ser Glu Leu Glu Glu Ile Glu Gln Ala Leu Lys
    210                 215                 220

agc tcg ctc aac gtg ctg gcc ccg ggt ggg cgg ctt tcg atc atc agc     720
Ser Ser Leu Asn Val Leu Ala Pro Gly Gly Arg Leu Ser Ile Ile Ser
225                 230                 235                 240

ttc cac tcg ctg gaa gac cgt att gtg aaa cgt ttt atg cgt gaa aac     768
Phe His Ser Leu Glu Asp Arg Ile Val Lys Arg Phe Met Arg Glu Asn
                245                 250                 255

agc cgc ggt ccg caa gtt ccg gca ggg tta ccg atg act gaa gag cag     816
Ser Arg Gly Pro Gln Val Pro Ala Gly Leu Pro Met Thr Glu Glu Gln
            260                 265                 270

ctc aaa aaa ctg ggt ggc cgt cag ctg cga gca cta ggc aag tta atg     864
Leu Lys Lys Leu Gly Gly Arg Gln Leu Arg Ala Leu Gly Lys Leu Met
        275                 280                 285

ccg ggc gaa gaa gag gtg gct gag aac cct cgt gcc cgt agt tca gtt     912
Pro Gly Glu Glu Glu Val Ala Glu Asn Pro Arg Ala Arg Ser Ser Val
    290                 295                 300

ctg cgt att gca gag agg acg aat gca tga                             942
Leu Arg Ile Ala Glu Arg Thr Asn Ala  *
305                 310
```

<210> 254
<211> 366
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(366)

<400> 254

```
atg atc agc aga gtg aca gaa gct cta agc aaa gtt aaa gga tcg atg     48
Met Ile Ser Arg Val Thr Glu Ala Leu Ser Lys Val Lys Gly Ser Met
 1               5                  10                  15

gga agc cac gag cgc cat gca ttg cct ggt gtt atc ggt gac gat ctt     96
Gly Ser His Glu Arg His Ala Leu Pro Gly Val Ile Gly Asp Asp Leu
            20                  25                  30

ttg cga ttt ggg aag ctg cca ctc tgc ctg ttc att tgc att att ttg     144
Leu Arg Phe Gly Lys Leu Pro Leu Cys Leu Phe Ile Cys Ile Ile Leu
```

35 40 45

acg gcg gtg act gtg gta acc acg gcg cac cat acc cgt tta ctg acc 192
Thr Ala Val Thr Val Val Thr Thr Ala His His Thr Arg Leu Leu Thr
50 55 60

gct cag cgc gaa caa ctg gtg ctg gag cga gat gct tta gac att gaa 240
Ala Gln Arg Glu Gln Leu Val Leu Glu Arg Asp Ala Leu Asp Ile Glu
65 70 75 80

tgg cgc aac ctg atc ctt gaa gag aat gcg ctc ggc gac cat agc cgg 288
Trp Arg Asn Leu Ile Leu Glu Glu Asn Ala Leu Gly Asp His Ser Arg
85 90 95

gtg gaa agg atc gcc acg gaa aag ctg caa atg cag cat gtt gat ccg 336
Val Glu Arg Ile Ala Thr Glu Lys Leu Gln Met Gln His Val Asp Pro
100 105 110

tca caa gaa aat atc gta gtg caa aaa taa 366
Ser Gln Glu Asn Ile Val Val Gln Lys *
115 120

<210> 255
<211> 1767
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1767)

<400> 255
atg aaa gca gcg gcg aaa acg cag aaa cca aaa cgt cag gaa gaa cat 48
Met Lys Ala Ala Ala Lys Thr Gln Lys Pro Lys Arg Gln Glu Glu His
1 5 10 15

gcc aac ttt atc agt tgg cgt ttt gcg ttg tta tgc ggc tgt att ctc 96
Ala Asn Phe Ile Ser Trp Arg Phe Ala Leu Leu Cys Gly Cys Ile Leu
20 25 30

ctg gcg ctg gct ttt ctg ctc gga cgc gta gcg tgg tta caa gtt atc 144
Leu Ala Leu Ala Phe Leu Leu Gly Arg Val Ala Trp Leu Gln Val Ile
35 40 45

tcc ccg gat atg ctg gtg aaa gag ggc gac atg cgt tct ctt cgc gtt 192
Ser Pro Asp Met Leu Val Lys Glu Gly Asp Met Arg Ser Leu Arg Val
50 55 60

cag caa gtt tcc acc tcc cgc ggc atg att act gac cgt tct ggt cgc 240
Gln Gln Val Ser Thr Ser Arg Gly Met Ile Thr Asp Arg Ser Gly Arg
65 70 75 80

ccg tta gcg gtg agc gtg ccg gta aaa gcg att tgg gct gac ccg aaa 288
Pro Leu Ala Val Ser Val Pro Val Lys Ala Ile Trp Ala Asp Pro Lys
85 90 95

gaa gtg cat gac gct ggc ggt atc agc gtc ggt gac cgc tgg aag gcg 336
Glu Val His Asp Ala Gly Gly Ile Ser Val Gly Asp Arg Trp Lys Ala
100 105 110

ctg gct aac gcg ctc aat att ccg ctg gat cag ctt tca gcc cgc att 384
Leu Ala Asn Ala Leu Asn Ile Pro Leu Asp Gln Leu Ser Ala Arg Ile
115 120 125

aac gcc aac ccg aaa ggg cgc ttt att tat ctg gcg cgt cag gtg aac 432
Asn Ala Asn Pro Lys Gly Arg Phe Ile Tyr Leu Ala Arg Gln Val Asn
130 135 140

cct gac atg gcg gac tac atc aaa aaa ctg aaa ctg ccg ggg att cat 480
Pro Asp Met Ala Asp Tyr Ile Lys Lys Leu Lys Leu Pro Gly Ile His
145 150 155 160

ctg cgt gaa gag tct cgc cgt tac tat ccg tcc ggc gaa gtg act gct 528
Leu Arg Glu Glu Ser Arg Arg Tyr Tyr Pro Ser Gly Glu Val Thr Ala
165 170 175

cac ctc atc ggc ttt act aac gtc gat agt caa ggg att gag ggc gtt 576
His Leu Ile Gly Phe Thr Asn Val Asp Ser Gln Gly Ile Glu Gly Val
180 185 190

gag aag agt ttc gat aaa tgg ctt acc ggg cag ccg ggt gag cgc att 624
Glu Lys Ser Phe Asp Lys Trp Leu Thr Gly Gln Pro Gly Glu Arg Ile
195 200 205

gtg cgt aaa gac cgc tat ggt cgc gta att gaa gat att tct tct act 672
Val Arg Lys Asp Arg Tyr Gly Arg Val Ile Glu Asp Ile Ser Ser Thr
210 215 220

gac agc cag gca gcg cac aac ctg gcg ctg agt att gat gaa cgc ctg 720
Asp Ser Gln Ala Ala His Asn Leu Ala Leu Ser Ile Asp Glu Arg Leu
225 230 235 240

cag gcg ctg gtt tat cgc gaa ctg aac aac gcg gtg gcc ttt aac aag 768
Gln Ala Leu Val Tyr Arg Glu Leu Asn Asn Ala Val Ala Phe Asn Lys
245 250 255

gct gaa tct ggt agc gcc gtg ctg gtg gat gtc aac acc ggt gaa gtg 816
Ala Glu Ser Gly Ser Ala Val Leu Val Asp Val Asn Thr Gly Glu Val
260 265 270

ctg gcg atg gct aac agc ccg tca tac aac cct aac aat ctg agc ggc 864
Leu Ala Met Ala Asn Ser Pro Ser Tyr Asn Pro Asn Asn Leu Ser Gly
275 280 285

acg ccg aaa gag gcg atg cgt aac cgt acc atc acc gac gtg ttt gaa 912
Thr Pro Lys Glu Ala Met Arg Asn Arg Thr Ile Thr Asp Val Phe Glu
290 295 300

ccg ggc tca acg gtt aaa ccg atg gtg gta atg acc gcg ttg caa cgt 960
Pro Gly Ser Thr Val Lys Pro Met Val Val Met Thr Ala Leu Gln Arg
305 310 315 320

ggc gtg gtg cgg gaa aac tcg gta ctc aat acc att cct tat cga att 1008
Gly Val Val Arg Glu Asn Ser Val Leu Asn Thr Ile Pro Tyr Arg Ile
325 330 335

aac ggc cac gaa atc aaa gac gtg gca cgc tac agc gaa tta acc ctg 1056
Asn Gly His Glu Ile Lys Asp Val Ala Arg Tyr Ser Glu Leu Thr Leu
340 345 350

acc ggg gta tta cag aag tcg agt aac gtc ggt gtt tcc aag ctg gcg 1104

Thr Gly Val Leu Gln Lys Ser Ser Asn Val Gly Val Ser Lys Leu Ala
355 360 365

tta gcg atg ccg tcc tca gcg tta gta gat act tac tca cgt ttt gga 1152
Leu Ala Met Pro Ser Ser Ala Leu Val Asp Thr Tyr Ser Arg Phe Gly
370 375 380

ctg gga aaa gcg acc aat ttg ggg ttg gtc gga gaa cgc agt ggc tta 1200
Leu Gly Lys Ala Thr Asn Leu Gly Leu Val Gly Glu Arg Ser Gly Leu
385 390 395 400

tat cct caa aaa caa cgg tgg tct gac ata gag agg gcc acc ttc tct 1248
Tyr Pro Gln Lys Gln Arg Trp Ser Asp Ile Glu Arg Ala Thr Phe Ser
405 410 415

ttc ggc tac ggg cta atg gta aca cca tta cag tta gcg cga gtc tac 1296
Phe Gly Tyr Gly Leu Met Val Thr Pro Leu Gln Leu Ala Arg Val Tyr
420 425 430

gca act atc ggc agc tac ggc att tat cgc cca ctg tcg att acc aaa 1344
Ala Thr Ile Gly Ser Tyr Gly Ile Tyr Arg Pro Leu Ser Ile Thr Lys
435 440 445

gtt gac ccc ccg gtt ccc ggt gaa cgt gtc ttc ccg gaa tcc att gtc 1392
Val Asp Pro Pro Val Pro Gly Glu Arg Val Phe Pro Glu Ser Ile Val
450 455 460

cgc act gtg gtg cat atg atg gaa agc gtg gcg cta cca ggc ggc ggc 1440
Arg Thr Val Val His Met Met Glu Ser Val Ala Leu Pro Gly Gly Gly
465 470 475 480

ggc gtg aag gcg gcg att aaa ggc tat cgt atc gcc att aaa acc ggt 1488
Gly Val Lys Ala Ala Ile Lys Gly Tyr Arg Ile Ala Ile Lys Thr Gly
485 490 495

acc gcg aaa aag gtc ggg ccg gac ggt cgc tac atc aat aaa tat att 1536
Thr Ala Lys Lys Val Gly Pro Asp Gly Arg Tyr Ile Asn Lys Tyr Ile
500 505 510

gct tat acc gca ggc gtt gcg cct gcg agt cag ccg cgc ttc gcg ctg 1584
Ala Tyr Thr Ala Gly Val Ala Pro Ala Ser Gln Pro Arg Phe Ala Leu
515 520 525

gtt gtt gtt atc aac gat ccg cag gcg ggt aaa tac tac ggc ggc gcc 1632
Val Val Val Ile Asn Asp Pro Gln Ala Gly Lys Tyr Tyr Gly Gly Ala
530 535 540

gtt tcc gcg ccg gtc ttt ggt gcc atc atg ggc ggc gta ttg cgt acc 1680
Val Ser Ala Pro Val Phe Gly Ala Ile Met Gly Gly Val Leu Arg Thr
545 550 555 560

atg aac atc gag ccg gat gcg ctg aca acg ggc gat aaa aat gaa ttt 1728
Met Asn Ile Glu Pro Asp Ala Leu Thr Thr Gly Asp Lys Asn Glu Phe
565 570 575

gtg att aat caa ggc gag ggg aca ggt ggc aga tcg taa 1767
Val Ile Asn Gln Gly Glu Gly Thr Gly Gly Arg Ser *
580 585

<210> 256

<211> 1488
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1488)

<400> 256
gtg gca gat cgt aat ttg cgc gac ctt ctt gct ccg tgg gtg cca gac 48
Met Ala Asp Arg Asn Leu Arg Asp Leu Leu Ala Pro Trp Val Pro Asp
1 5 10 15

gca cct tcg cga gca ctg cga gag atg aca ctc gac agc cgt gtg gct 96
Ala Pro Ser Arg Ala Leu Arg Glu Met Thr Leu Asp Ser Arg Val Ala
20 25 30

gcg gcg ggc gat ctc ttt gta gct gta gta ggt cat cag gcg gac ggg 144
Ala Ala Gly Asp Leu Phe Val Ala Val Val Gly His Gln Ala Asp Gly
35 40 45

cgt cga tat atc ccg cag gcg ata gcg caa ggt gtg gct gcc att att 192
Arg Arg Tyr Ile Pro Gln Ala Ile Ala Gln Gly Val Ala Ala Ile Ile
50 55 60

gca gag gcg aaa gat gag gcg acc gat ggt gaa atc cgt gaa atg cac 240
Ala Glu Ala Lys Asp Glu Ala Thr Asp Gly Glu Ile Arg Glu Met His
65 70 75 80

ggc gta ccg gtc atc tat ctc agc cag ctc aac gag cgt tta tct gca 288
Gly Val Pro Val Ile Tyr Leu Ser Gln Leu Asn Glu Arg Leu Ser Ala
85 90 95

ctg gcg ggc cgc ttt tac cat gaa ccc tct gac aat tta cgt ctc gtg 336
Leu Ala Gly Arg Phe Tyr His Glu Pro Ser Asp Asn Leu Arg Leu Val
100 105 110

ggc gta acg ggc acc aac ggc aaa acc acg act acc cag ctg ttg gcg 384
Gly Val Thr Gly Thr Asn Gly Lys Thr Thr Thr Thr Gln Leu Leu Ala
115 120 125

cag tgg agc caa ctg ctt ggc gaa atc agc gcg gta atg ggc acc gtt 432
Gln Trp Ser Gln Leu Leu Gly Glu Ile Ser Ala Val Met Gly Thr Val
130 135 140

ggt aac ggc ctg ctg ggg aaa gtg atc ccg aca gaa aat aca acc ggt 480
Gly Asn Gly Leu Leu Gly Lys Val Ile Pro Thr Glu Asn Thr Thr Gly
145 150 155 160

tcg gca gtc gat gtt cag cat gag ctg gcg ggg ctg gtg gat cag ggc 528
Ser Ala Val Asp Val Gln His Glu Leu Ala Gly Leu Val Asp Gln Gly
165 170 175

gcg acg ttt tgc gca atg gaa gtt tcc tcc cac ggg ctg gta cag cac 576
Ala Thr Phe Cys Ala Met Glu Val Ser Ser His Gly Leu Val Gln His
180 185 190

cgt gtg gcg gca ttg aaa ttt gcg gcg tcg gtc ttt acc aac tta agc 624
Arg Val Ala Ala Leu Lys Phe Ala Ala Ser Val Phe Thr Asn Leu Ser
195 200 205

```
cgc gat cac ctt gat tat cat ggt gat atg gaa cac tac gaa gcc gcg     672
Arg Asp His Leu Asp Tyr His Gly Asp Met Glu His Tyr Glu Ala Ala
    210                 215                 220

aaa tgg ctg ctt tat tct gag cat cat tgc ggt cag gcg att att aac     720
Lys Trp Leu Leu Tyr Ser Glu His His Cys Gly Gln Ala Ile Ile Asn
225                 230                 235                 240

gcc gac gat gaa gtg ggc cgc cgc tgg ctg gca aaa ctg ccg gac gcg     768
Ala Asp Asp Glu Val Gly Arg Arg Trp Leu Ala Lys Leu Pro Asp Ala
                245                 250                 255

gtt gcg gta tca atg gaa gat cat att aat ccg aac tgt cac gga cgc     816
Val Ala Val Ser Met Glu Asp His Ile Asn Pro Asn Cys His Gly Arg
            260                 265                 270

tgg ttg aaa gcg acc gaa gtg aac tat cac gac agc ggt gcg acg att     864
Trp Leu Lys Ala Thr Glu Val Asn Tyr His Asp Ser Gly Ala Thr Ile
        275                 280                 285

cgc ttt agc tca agt tgg ggc gat ggc gaa att gaa agc cat ctg atg     912
Arg Phe Ser Ser Ser Trp Gly Asp Gly Glu Ile Glu Ser His Leu Met
    290                 295                 300

ggc gct ttt aac gtc agc aac ctg ctg ctc gcg ctg gcg aca ctg ttg     960
Gly Ala Phe Asn Val Ser Asn Leu Leu Leu Ala Leu Ala Thr Leu Leu
305                 310                 315                 320

gca ctc ggc tat cca ctg gct gat ctg ctg aaa acc gcc gcg cgt ctg     1008
Ala Leu Gly Tyr Pro Leu Ala Asp Leu Leu Lys Thr Ala Ala Arg Leu
                325                 330                 335

caa ccg gtt tgc gga cgt atg gaa gtg ttc act gcg cca ggc aaa ccg     1056
Gln Pro Val Cys Gly Arg Met Glu Val Phe Thr Ala Pro Gly Lys Pro
            340                 345                 350

acg gtg gtg gtg gat tac gcg cat acg ccg gat gca ctg gaa aaa gcc     1104
Thr Val Val Val Asp Tyr Ala His Thr Pro Asp Ala Leu Glu Lys Ala
        355                 360                 365

tta cag gcg gcg cgt ctg cac tgt gcg ggc aag ctg tgg tgt gtc ttt     1152
Leu Gln Ala Ala Arg Leu His Cys Ala Gly Lys Leu Trp Cys Val Phe
    370                 375                 380

ggc tgt ggt ggc gat cgc gat aaa ggt aag cgt cca ctg atg ggc gca     1200
Gly Cys Gly Gly Asp Arg Asp Lys Gly Lys Arg Pro Leu Met Gly Ala
385                 390                 395                 400

att gcc gaa gag ttt gct gac gtg gcg gtg gtg acg gac gat aac ccg     1248
Ile Ala Glu Glu Phe Ala Asp Val Ala Val Val Thr Asp Asp Asn Pro
                405                 410                 415

cgt acc gaa gaa ccg cgt gcc atc atc aac gat att ctg gcg gga atg     1296
Arg Thr Glu Glu Pro Arg Ala Ile Ile Asn Asp Ile Leu Ala Gly Met
            420                 425                 430

tta gat gcc gga cat gcc aaa gtg atg gaa ggc cgt gct gaa gcg gtg     1344
Leu Asp Ala Gly His Ala Lys Val Met Glu Gly Arg Ala Glu Ala Val
        435                 440                 445

act tgc gcc gtt atg cag gct aaa gag aat gat gtg gta ctg gtc gcg     1392
```

```
Thr Cys Ala Val Met Gln Ala Lys Glu Asn Asp Val Val Leu Val Ala
    450                 455                 460

ggc aaa ggc cat gaa gat tac cag att gtt ggc aat cag cgt ctg gac    1440
Gly Lys Gly His Glu Asp Tyr Gln Ile Val Gly Asn Gln Arg Leu Asp
465                 470                 475                 480

tac tcc gat cgc gtc acg gtg gcg cgt ctg ctg ggg gtg att gca tga    1488
Tyr Ser Asp Arg Val Thr Val Ala Arg Leu Leu Gly Val Ile Ala  *
                485                 490                 495
```

<210> 257
<211> 1359
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1359)

<400> 257

```
atg att agc gta acc ctt agc caa ctt acc gac att ctc aac ggt gaa    48
Met Ile Ser Val Thr Leu Ser Gln Leu Thr Asp Ile Leu Asn Gly Glu
 1               5                  10                  15

ctg caa ggt gca gat atc acc ctt gat gct gta acc act gat acc cga    96
Leu Gln Gly Ala Asp Ile Thr Leu Asp Ala Val Thr Thr Asp Thr Arg
            20                  25                  30

aaa ctg acg ccg ggc tgc ctg ttt gtt gcc ctg aaa ggc gaa cgt ttt    144
Lys Leu Thr Pro Gly Cys Leu Phe Val Ala Leu Lys Gly Glu Arg Phe
        35                  40                  45

gat gcc cac gat ttt gcc gac cag gcg aaa gct ggc ggc gca ggc gca    192
Asp Ala His Asp Phe Ala Asp Gln Ala Lys Ala Gly Gly Ala Gly Ala
    50                  55                  60

cta ctg gtt agc cgt ccg ctg gac atc gac ctg ccg cag tta atc gtc    240
Leu Leu Val Ser Arg Pro Leu Asp Ile Asp Leu Pro Gln Leu Ile Val
65                  70                  75                  80

aag gat acg cgt ctg gcg ttt ggt gaa ctg gct gca tgg gtt cgc cag    288
Lys Asp Thr Arg Leu Ala Phe Gly Glu Leu Ala Ala Trp Val Arg Gln
                85                  90                  95

caa gtt ccg gcg cgc gtg gtt gct ctg acg ggg tcc tcc ggc aaa acc    336
Gln Val Pro Ala Arg Val Val Ala Leu Thr Gly Ser Ser Gly Lys Thr
            100                 105                 110

tcc gtt aaa gag atg acg gcg gcg att tta agc cag tgc ggc aac acg    384
Ser Val Lys Glu Met Thr Ala Ala Ile Leu Ser Gln Cys Gly Asn Thr
        115                 120                 125

ctt tat acg gca ggc aat ctc aac aac gac atc ggt gta ccg atg acg    432
Leu Tyr Thr Ala Gly Asn Leu Asn Asn Asp Ile Gly Val Pro Met Thr
    130                 135                 140

ctg ttg cgc tta acg ccg gaa tac gat tac gca gtt att gaa ctt ggc    480
Leu Leu Arg Leu Thr Pro Glu Tyr Asp Tyr Ala Val Ile Glu Leu Gly
```

145 150 155 160

gcg aac cat cag ggc gaa ata gcc tgg act gtg agt ctg act cgc ccg 528
Ala Asn His Gln Gly Glu Ile Ala Trp Thr Val Ser Leu Thr Arg Pro
165 170 175

gaa gct gcg ctg gtc aac aac ctg gca gcg gcg cat ctg gaa ggt ttt 576
Glu Ala Ala Leu Val Asn Asn Leu Ala Ala Ala His Leu Glu Gly Phe
180 185 190

ggc tcg ctt gcg ggt gtc gcg aaa gcg aaa ggt gaa atc ttt agc ggc 624
Gly Ser Leu Ala Gly Val Ala Lys Ala Lys Gly Glu Ile Phe Ser Gly
195 200 205

ctg ccg gaa aac ggt atc gcc att atg aac gcc gac aac aac gac tgg 672
Leu Pro Glu Asn Gly Ile Ala Ile Met Asn Ala Asp Asn Asn Asp Trp
210 215 220

ctg aac tgg cag agc gta att ggc tca cgc aaa gtg tgg cgt ttc tca 720
Leu Asn Trp Gln Ser Val Ile Gly Ser Arg Lys Val Trp Arg Phe Ser
225 230 235 240

ccc aat gcc gcc aac agc gat ttc acc gcc acc aat atc cat gtg acc 768
Pro Asn Ala Ala Asn Ser Asp Phe Thr Ala Thr Asn Ile His Val Thr
245 250 255

tcg cac ggt acg gaa ttt acc cta caa acc cca acc ggt agc gtc gat 816
Ser His Gly Thr Glu Phe Thr Leu Gln Thr Pro Thr Gly Ser Val Asp
260 265 270

gtt ctg ctg ccg ttg ccg ggg cgt cac aat att gcg aat gcg ctg gca 864
Val Leu Leu Pro Leu Pro Gly Arg His Asn Ile Ala Asn Ala Leu Ala
275 280 285

gcc gct gcg ctc tcc atg tcc gtg ggc gca acg ctt gat gct atc aaa 912
Ala Ala Ala Leu Ser Met Ser Val Gly Ala Thr Leu Asp Ala Ile Lys
290 295 300

gcg ggg ctg gca aat ctg aaa gct gtt cca ggc cgt ctg ttc ccc atc 960
Ala Gly Leu Ala Asn Leu Lys Ala Val Pro Gly Arg Leu Phe Pro Ile
305 310 315 320

caa ctg gca gaa aac cag ttg ctg ctc gac gac tcc tac aac gcc aat 1008
Gln Leu Ala Glu Asn Gln Leu Leu Leu Asp Asp Ser Tyr Asn Ala Asn
325 330 335

gtc ggt tca atg act gca gca gtc cag gta ctg gct gaa atg ccg ggc 1056
Val Gly Ser Met Thr Ala Ala Val Gln Val Leu Ala Glu Met Pro Gly
340 345 350

tac cgc gtg ctg gtg gtg ggc gat atg gcg gaa ctg ggc gct gaa agc 1104
Tyr Arg Val Leu Val Val Gly Asp Met Ala Glu Leu Gly Ala Glu Ser
355 360 365

gaa gcc tgc cat gta cag gtg ggc gag gcg gca aaa gct gct ggt att 1152
Glu Ala Cys His Val Gln Val Gly Glu Ala Ala Lys Ala Ala Gly Ile
370 375 380

gac cgc gtg tta agc gtg ggt aaa caa agc cat gct atc agc acc gcc 1200
Asp Arg Val Leu Ser Val Gly Lys Gln Ser His Ala Ile Ser Thr Ala
385 390 395 400

```
agc ggc gtt ggc gaa cat ttt gct gat aaa act gcg tta att acg cgt    1248
Ser Gly Val Gly Glu His Phe Ala Asp Lys Thr Ala Leu Ile Thr Arg
                405                 410                 415

ctt aaa tta ctg att gct gag caa cag gta att acg att tta gtt aag    1296
Leu Lys Leu Leu Ile Ala Glu Gln Gln Val Ile Thr Ile Leu Val Lys
            420                 425                 430

ggt tca cgt agt gcc gcc atg gaa gag gta gta cgc gct tta cag gag    1344
Gly Ser Arg Ser Ala Ala Met Glu Glu Val Val Arg Ala Leu Gln Glu
        435                 440                 445

aat ggg aca tgt tag                                                1359
Asn Gly Thr Cys  *
    450


<210> 258
<211> 1083
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1083)

<400> 258
atg tta gtt tgg ctg gcc gaa cat ttg gtc aaa tat tat tcc ggc ttt    48
Met Leu Val Trp Leu Ala Glu His Leu Val Lys Tyr Tyr Ser Gly Phe
 1               5                  10                  15

aac gtc ttt tcc tat ctg acg ttt cgc gcc atc gtc agc ctg ctg acc    96
Asn Val Phe Ser Tyr Leu Thr Phe Arg Ala Ile Val Ser Leu Leu Thr
            20                  25                  30

gcg ctg ttc atc tca ttg tgg atg ggc ccg cgt atg att gct cat ttg    144
Ala Leu Phe Ile Ser Leu Trp Met Gly Pro Arg Met Ile Ala His Leu
        35                  40                  45

caa aaa ctt tcc ttt ggt cag gtg gtg cgt aac gac ggt cct gaa tca    192
Gln Lys Leu Ser Phe Gly Gln Val Val Arg Asn Asp Gly Pro Glu Ser
    50                  55                  60

cac ttc agc aag cgc ggt acg ccg acc atg ggc ggg att atg atc ctg    240
His Phe Ser Lys Arg Gly Thr Pro Thr Met Gly Gly Ile Met Ile Leu
65                  70                  75                  80

acg gcg att gtg atc tcc gta ctg ctg tgg gct tac ccg tcc aat ccg    288
Thr Ala Ile Val Ile Ser Val Leu Leu Trp Ala Tyr Pro Ser Asn Pro
                85                  90                  95

tac gtc tgg tgc gtg ttg gtg gtg ctg gta ggt tac ggt gtt att ggc    336
Tyr Val Trp Cys Val Leu Val Val Leu Val Gly Tyr Gly Val Ile Gly
            100                 105                 110

ttt gtt gat gat tat cgc aaa gtg gtg cgt aaa gac acc aaa ggg ttg    384
Phe Val Asp Asp Tyr Arg Lys Val Val Arg Lys Asp Thr Lys Gly Leu
        115                 120                 125

atc gct cgt tgg aag tat ttc tgg atg tcg gtc att gcg ctg ggt gtc    432
```

```
Ile Ala Arg Trp Lys Tyr Phe Trp Met Ser Val Ile Ala Leu Gly Val
    130                 135                 140

gcc ttc gcc ctg tac ctt gcc ggc aaa gac acg ccc gca acg cag ctg      480
Ala Phe Ala Leu Tyr Leu Ala Gly Lys Asp Thr Pro Ala Thr Gln Leu
145                 150                 155                 160

gtg gtc cca ttc ttt aaa gat gtg atg ccg cag ctg ggg ctg ttc tac      528
Val Val Pro Phe Phe Lys Asp Val Met Pro Gln Leu Gly Leu Phe Tyr
                165                 170                 175

att ctg ctg gct tac ttc gtc att gtg ggt act ggc aac gcg gta aac      576
Ile Leu Leu Ala Tyr Phe Val Ile Val Gly Thr Gly Asn Ala Val Asn
            180                 185                 190

ctg acc gat ggt ctc gac ggc ctg gca att atg ccg acc gta ttt gtc      624
Leu Thr Asp Gly Leu Asp Gly Leu Ala Ile Met Pro Thr Val Phe Val
        195                 200                 205

gcc ggt ggt ttt gcg ctg gtg gcg tgg gcg acc ggc aat atg aac ttt      672
Ala Gly Gly Phe Ala Leu Val Ala Trp Ala Thr Gly Asn Met Asn Phe
    210                 215                 220

gcc agc tac ttg cat ata ccg tat ctg cga cac gcc ggg gaa ctg gtt      720
Ala Ser Tyr Leu His Ile Pro Tyr Leu Arg His Ala Gly Glu Leu Val
225                 230                 235                 240

att gtc tgt acc gcg ata gtc ggg gca gga ctg ggc ttc ctg tgg ttt      768
Ile Val Cys Thr Ala Ile Val Gly Ala Gly Leu Gly Phe Leu Trp Phe
                245                 250                 255

aac acc tat ccg gcg cag gtc ttt atg ggc gat gta ggt tcg ctg gcg      816
Asn Thr Tyr Pro Ala Gln Val Phe Met Gly Asp Val Gly Ser Leu Ala
            260                 265                 270

tta ggt ggt gcg tta ggc att atc gcc gta ctg cta cgt cag gaa ttc      864
Leu Gly Gly Ala Leu Gly Ile Ile Ala Val Leu Leu Arg Gln Glu Phe
        275                 280                 285

ctg ctg gtg att atg ggg ggc gtg ttc gtg gta gaa acg ctt tct gtc      912
Leu Leu Val Ile Met Gly Gly Val Phe Val Val Glu Thr Leu Ser Val
    290                 295                 300

atc ctg cag gtc ggc tcc ttt aaa ctg cgc gga caa cgt att ttc cgc      960
Ile Leu Gln Val Gly Ser Phe Lys Leu Arg Gly Gln Arg Ile Phe Arg
305                 310                 315                 320

atg gca ccg att cat cac cac tat gaa ctg aaa ggc tgg ccg gaa ccg      1008
Met Ala Pro Ile His His His Tyr Glu Leu Lys Gly Trp Pro Glu Pro
                325                 330                 335

cgc gtc att gtg cgt ttc tgg att att tcg ctg atg ctg gtt ctg att      1056
Arg Val Ile Val Arg Phe Trp Ile Ile Ser Leu Met Leu Val Leu Ile
            340                 345                 350

ggt ctg gca acg ctg aag gta cgt taa                                  1083
Gly Leu Ala Thr Leu Lys Val Arg  *
        355                 360
```

<210> 259

<211> 1317
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1317)

<400> 259
atg gct gat tat cag ggt aaa aat gtc gtc att atc ggc ctg ggc ctc 48
Met Ala Asp Tyr Gln Gly Lys Asn Val Val Ile Ile Gly Leu Gly Leu
1 5 10 15

acc ggg ctt tcc tgc gtg gac ttt ttc ctc gct cgc ggt gtg acg ccg 96
Thr Gly Leu Ser Cys Val Asp Phe Phe Leu Ala Arg Gly Val Thr Pro
20 25 30

cgc gtt atg gat acg cgt atg aca ccg cct ggc ctg gat aaa tta ccc 144
Arg Val Met Asp Thr Arg Met Thr Pro Pro Gly Leu Asp Lys Leu Pro
35 40 45

gaa gcc gta gaa cgc cac acg ggc agt ctg aat gat gaa tgg ctg atg 192
Glu Ala Val Glu Arg His Thr Gly Ser Leu Asn Asp Glu Trp Leu Met
50 55 60

gcg gca gat ctg att gtc gcc agt ccc ggt att gca ctg gcg cat cca 240
Ala Ala Asp Leu Ile Val Ala Ser Pro Gly Ile Ala Leu Ala His Pro
65 70 75 80

tcc tta agc gct gcc gct gat gcc gga atc gaa atc gtt ggc gat atc 288
Ser Leu Ser Ala Ala Ala Asp Ala Gly Ile Glu Ile Val Gly Asp Ile
85 90 95

gag ctg ttc tgt cgc gaa gca caa gca ccg att gtg gcg att acc ggt 336
Glu Leu Phe Cys Arg Glu Ala Gln Ala Pro Ile Val Ala Ile Thr Gly
100 105 110

tct aac ggc aaa agc acg gtc acc acg cta gtg ggt gaa atg gcg aaa 384
Ser Asn Gly Lys Ser Thr Val Thr Thr Leu Val Gly Glu Met Ala Lys
115 120 125

gcg gcg ggg gtt aac gtt ggt gtg ggt ggc aat att ggc ctg cct gcg 432
Ala Ala Gly Val Asn Val Gly Val Gly Gly Asn Ile Gly Leu Pro Ala
130 135 140

ttg atg cta ctg gat gat gag tgt gaa ctg tac gtg ctg gaa ctg tcg 480
Leu Met Leu Leu Asp Asp Glu Cys Glu Leu Tyr Val Leu Glu Leu Ser
145 150 155 160

agc ttc cag ctg gaa acc acc tcc agc tta cag gcg gta gca gcg acc 528
Ser Phe Gln Leu Glu Thr Thr Ser Ser Leu Gln Ala Val Ala Ala Thr
165 170 175

att ctg aac gtg act gaa gat cat atg gat cgc tat ccg ttt ggt tta 576
Ile Leu Asn Val Thr Glu Asp His Met Asp Arg Tyr Pro Phe Gly Leu
180 185 190

caa cag tat cgt gca gca aaa ctg cgc att tac gaa aac gcg aaa gtt 624
Gln Gln Tyr Arg Ala Ala Lys Leu Arg Ile Tyr Glu Asn Ala Lys Val
195 200 205

```
tgc gtg gtt aat gct gat gat gcc tta aca atg ccg att cgc ggt gcg     672
Cys Val Val Asn Ala Asp Asp Ala Leu Thr Met Pro Ile Arg Gly Ala
    210                 215                 220

gat gaa cgc tgc gtc agc ttt ggc gtc aac atg ggt gac tat cac ctg     720
Asp Glu Arg Cys Val Ser Phe Gly Val Asn Met Gly Asp Tyr His Leu
225                 230                 235                 240

aat cat cag cag ggc gaa acc tgg ctg cgg gtt aaa ggc gag aaa gtg     768
Asn His Gln Gln Gly Glu Thr Trp Leu Arg Val Lys Gly Glu Lys Val
                245                 250                 255

ctg aat gtg aaa gag atg aaa ctt tcc ggg cag cat aac tac acc aat     816
Leu Asn Val Lys Glu Met Lys Leu Ser Gly Gln His Asn Tyr Thr Asn
            260                 265                 270

gcg ctg gcg gcg ctg gcg ctg gca gat gct gca ggg tta ccg cgt gcc     864
Ala Leu Ala Ala Leu Ala Leu Ala Asp Ala Ala Gly Leu Pro Arg Ala
        275                 280                 285

agc agc ctg aaa gcg tta acc aca ttc act ggt ctg ccg cat cgc ttt     912
Ser Ser Leu Lys Ala Leu Thr Thr Phe Thr Gly Leu Pro His Arg Phe
    290                 295                 300

gaa gtt gtg ctg gag cat aac ggc gta cgt tgg att aac gat tcg aaa     960
Glu Val Val Leu Glu His Asn Gly Val Arg Trp Ile Asn Asp Ser Lys
305                 310                 315                 320

gcg acc aac gtc ggc agt acg gaa gcg gcg ctg aat ggc ctg cac gta    1008
Ala Thr Asn Val Gly Ser Thr Glu Ala Ala Leu Asn Gly Leu His Val
                325                 330                 335

gac ggc aca ctg cat ttg ttg ctg ggt ggc gat ggt aaa tcg gcg gac    1056
Asp Gly Thr Leu His Leu Leu Leu Gly Gly Asp Gly Lys Ser Ala Asp
            340                 345                 350

ttt agc cca ctg gcg cgt tac ctg aat ggc gat aac gta cgt ctg tat    1104
Phe Ser Pro Leu Ala Arg Tyr Leu Asn Gly Asp Asn Val Arg Leu Tyr
        355                 360                 365

tgt ttc ggt cgt gac ggc gcg cag ctg gcg gcg cta cgc ccg gaa gtg    1152
Cys Phe Gly Arg Asp Gly Ala Gln Leu Ala Ala Leu Arg Pro Glu Val
    370                 375                 380

gca gaa caa acc gaa act atg gaa cag gcg atg cgc ttg ctg gct ccg    1200
Ala Glu Gln Thr Glu Thr Met Glu Gln Ala Met Arg Leu Leu Ala Pro
385                 390                 395                 400

cgt gtt cag ccg ggc gat atg gtt ctg ctc tcc cca gcc tgt gcc agc    1248
Arg Val Gln Pro Gly Asp Met Val Leu Leu Ser Pro Ala Cys Ala Ser
                405                 410                 415

ctt gat cag ttc aag aac ttt gaa caa cga ggc aat gag ttt gcc cgt    1296
Leu Asp Gln Phe Lys Asn Phe Glu Gln Arg Gly Asn Glu Phe Ala Arg
            420                 425                 430

ctg gcg aag gag tta ggt tga                                        1317
Leu Ala Lys Glu Leu Gly  *
        435
```

<210> 260
<211> 1245
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1245)

<400> 260
```
atg cgt tta tct ctc cct cgc ctg aaa atg ccg cgc ctg cca gga ttc    48
Met Arg Leu Ser Leu Pro Arg Leu Lys Met Pro Arg Leu Pro Gly Phe
 1               5                  10                  15

agt atc ctg gtc tgg atc tcc acg gcg cta aag ggc tgg gtg atg ggc    96
Ser Ile Leu Val Trp Ile Ser Thr Ala Leu Lys Gly Trp Val Met Gly
            20                  25                  30

tcg cgg gaa aaa gat acc gac agc ctg atc atg tac gat cgc acc tta   144
Ser Arg Glu Lys Asp Thr Asp Ser Leu Ile Met Tyr Asp Arg Thr Leu
        35                  40                  45

ctg tgg ctg acc ttc ggc ctc gcg gcg att ggc ttt atc atg gtg acc   192
Leu Trp Leu Thr Phe Gly Leu Ala Ala Ile Gly Phe Ile Met Val Thr
    50                  55                  60

tcg gcg tca atg ccc ata ggg caa cgc tta acc aac gat ccg ttc ttc   240
Ser Ala Ser Met Pro Ile Gly Gln Arg Leu Thr Asn Asp Pro Phe Phe
 65                  70                  75                  80

ttc gcg aag cgt gat ggt gtc tat ctg att ttg gcg ttt att ctg gcg   288
Phe Ala Lys Arg Asp Gly Val Tyr Leu Ile Leu Ala Phe Ile Leu Ala
                 85                  90                  95

atc att acg ctg cgt ctg ccg atg gag ttc tgg caa cgc tac agt gcc   336
Ile Ile Thr Leu Arg Leu Pro Met Glu Phe Trp Gln Arg Tyr Ser Ala
            100                 105                 110

acg atg ctg ctc gga tct atc atc ctg ctg atg atc gtc ctg gta gtg   384
Thr Met Leu Leu Gly Ser Ile Ile Leu Leu Met Ile Val Leu Val Val
        115                 120                 125

ggt agc tcg gtt aaa ggg gca tcg cgt tgg atc gat ctc ggt ttg ctg   432
Gly Ser Ser Val Lys Gly Ala Ser Arg Trp Ile Asp Leu Gly Leu Leu
    130                 135                 140

cgt atc cag cct gcg gag ctg aca aaa ctg tcg ctg ttt tgc tat atc   480
Arg Ile Gln Pro Ala Glu Leu Thr Lys Leu Ser Leu Phe Cys Tyr Ile
145                 150                 155                 160

gcc aac tat ctg gtg cgt aaa ggc gac gaa gta cgt aat aac ctg cgc   528
Ala Asn Tyr Leu Val Arg Lys Gly Asp Glu Val Arg Asn Asn Leu Arg
                165                 170                 175

ggc ttc ctg aaa ccg atg ggc gtg att ctg gtg ttg gca gtg tta ctg   576
Gly Phe Leu Lys Pro Met Gly Val Ile Leu Val Leu Ala Val Leu Leu
            180                 185                 190

ctg gca cag cca gac ctt ggt acg gtg gtg gtg ttg ttt gtg act acg   624
Leu Ala Gln Pro Asp Leu Gly Thr Val Val Val Leu Phe Val Thr Thr
        195                 200                 205
```

```
ctg gcg atg ttg ttc ctg gcg gga gcg aaa ttg tgg cag ttc att gcc    672
Leu Ala Met Leu Phe Leu Ala Gly Ala Lys Leu Trp Gln Phe Ile Ala
    210                 215                 220

att atc ggt atg ggc att tca gcg gtt gtg ttg ctg ata ctc gcc gaa    720
Ile Ile Gly Met Gly Ile Ser Ala Val Val Leu Leu Ile Leu Ala Glu
225                 230                 235                 240

ccg tac cgt atc cgc cgt gtt acc gca ttc tgg aac ccg tgg gaa gat    768
Pro Tyr Arg Ile Arg Arg Val Thr Ala Phe Trp Asn Pro Trp Glu Asp
                245                 250                 255

ccc ttt ggc agc ggc tat cag tta acg caa tcg ctg atg gcg ttt ggt    816
Pro Phe Gly Ser Gly Tyr Gln Leu Thr Gln Ser Leu Met Ala Phe Gly
            260                 265                 270

cgc ggc gaa ctt tgg ggg caa ggt tta ggt aac tcg gta caa aaa ctg    864
Arg Gly Glu Leu Trp Gly Gln Gly Leu Gly Asn Ser Val Gln Lys Leu
        275                 280                 285

gag tat ctg ccg gaa gcg cac act gac ttt att ttc gcc att atc ggc    912
Glu Tyr Leu Pro Glu Ala His Thr Asp Phe Ile Phe Ala Ile Ile Gly
    290                 295                 300

gaa gaa ctg ggg tat gtc ggt gtg gtg ctg gca ctt tta atg gta ttc    960
Glu Glu Leu Gly Tyr Val Gly Val Val Leu Ala Leu Leu Met Val Phe
305                 310                 315                 320

ttc gtc gct ttt cgc gcg atg tcg att ggc cgt aaa gca tta gaa att    1008
Phe Val Ala Phe Arg Ala Met Ser Ile Gly Arg Lys Ala Leu Glu Ile
                325                 330                 335

gac cac cgt ttt tcc ggt ttt ctc gcc tgt tct att ggc atc tgg ttt    1056
Asp His Arg Phe Ser Gly Phe Leu Ala Cys Ser Ile Gly Ile Trp Phe
            340                 345                 350

agc ttc cag gcg ctg gtt aac gta ggc gcg gcg gcg ggg atg tta ccg    1104
Ser Phe Gln Ala Leu Val Asn Val Gly Ala Ala Ala Gly Met Leu Pro
        355                 360                 365

acc aaa ggt ctg aca ttg ccg ctg atc agt tac ggt ggt tcg agc tta    1152
Thr Lys Gly Leu Thr Leu Pro Leu Ile Ser Tyr Gly Gly Ser Ser Leu
    370                 375                 380

ctg att atg tcg aca gcc atc atg atg ctg ttg cgt att gat tat gaa    1200
Leu Ile Met Ser Thr Ala Ile Met Met Leu Leu Arg Ile Asp Tyr Glu
385                 390                 395                 400

acg cgt ctg gag aaa gcg cag gcg ttt gta cga ggt tca cga tga        1245
Thr Arg Leu Glu Lys Ala Gln Ala Phe Val Arg Gly Ser Arg  *
                405                 410
```

<210> 261
<211> 1068
<212> DNA
<213> Escherichia coli

<220>

```
<221> CDS
<222> (1)...(1068)

<400> 261
atg agt ggt caa gga aag cga tta atg gtg atg gca ggc gga acc ggt    48
Met Ser Gly Gln Gly Lys Arg Leu Met Val Met Ala Gly Gly Thr Gly
 1               5                  10                  15

gga cat gta ttc ccg gga ctg gcg gtt gcg cac cat cta atg gct cag    96
Gly His Val Phe Pro Gly Leu Ala Val Ala His His Leu Met Ala Gln
            20                  25                  30

ggt tgg caa gtt cgc tgg ctg ggg act gcc gac cgt atg gaa gcg gac   144
Gly Trp Gln Val Arg Trp Leu Gly Thr Ala Asp Arg Met Glu Ala Asp
        35                  40                  45

tta gtg cca aaa cat ggc atc gaa att gat ttc att cgt atc tct ggt   192
Leu Val Pro Lys His Gly Ile Glu Ile Asp Phe Ile Arg Ile Ser Gly
    50                  55                  60

ctg cgt gga aaa ggt ata aaa gca ctg ata gct gcc ccg ctg cgt atc   240
Leu Arg Gly Lys Gly Ile Lys Ala Leu Ile Ala Ala Pro Leu Arg Ile
65                  70                  75                  80

ttc aac gcc tgg cgt cag gcg cgg gcg att atg aaa gcg tac aaa cct   288
Phe Asn Ala Trp Arg Gln Ala Arg Ala Ile Met Lys Ala Tyr Lys Pro
                85                  90                  95

gac gtg gtg ctc ggt atg gga ggc tac gtg tca ggt cca ggt ggt ctg   336
Asp Val Val Leu Gly Met Gly Gly Tyr Val Ser Gly Pro Gly Gly Leu
            100                 105                 110

gcc gcg tgg tcg tta ggc att ccg gtt gta ctt cat gaa caa aac ggt   384
Ala Ala Trp Ser Leu Gly Ile Pro Val Val Leu His Glu Gln Asn Gly
        115                 120                 125

att gcg ggc tta acc aat aaa tgg ctg gcg aag att gcc acc aaa gtg   432
Ile Ala Gly Leu Thr Asn Lys Trp Leu Ala Lys Ile Ala Thr Lys Val
    130                 135                 140

atg cag gcg ttt cca ggt gct ttc cct aat gcg gaa gta gtg ggt aac   480
Met Gln Ala Phe Pro Gly Ala Phe Pro Asn Ala Glu Val Val Gly Asn
145                 150                 155                 160

ccg gtg cgt acc gat gtg ttg gcg ctg ccg ttg ccg cag caa cgt ttg   528
Pro Val Arg Thr Asp Val Leu Ala Leu Pro Leu Pro Gln Gln Arg Leu
                165                 170                 175

gct gga cgt gaa ggt ccg gtt cgt gtg ctg gta gtg ggt ggt tct cag   576
Ala Gly Arg Glu Gly Pro Val Arg Val Leu Val Val Gly Gly Ser Gln
            180                 185                 190

ggc gca cgc att ctt aac cag aca atg ccg cag gtt gct gcg aaa ctg   624
Gly Ala Arg Ile Leu Asn Gln Thr Met Pro Gln Val Ala Ala Lys Leu
        195                 200                 205

ggt gat tca gtc act atc tgg cat cag agc ggc aaa ggt tcg caa caa   672
Gly Asp Ser Val Thr Ile Trp His Gln Ser Gly Lys Gly Ser Gln Gln
    210                 215                 220

tcc gtt gaa cag gcg tat gcc gaa gcg ggg caa ccg cag cat aaa gtg   720
```

Ser Val Glu Gln Ala Tyr Ala Glu Ala Gly Gln Pro Gln His Lys Val
225 230 235 240

acg gaa ttt att gat gat atg gcg gcg gcg tat gcg tgg gcg gat gtc 768
Thr Glu Phe Ile Asp Asp Met Ala Ala Ala Tyr Ala Trp Ala Asp Val
245 250 255

gtc gtt tgc cgc tcc ggt gcg tta acg gtg agt gaa atc gcc gcg gca 816
Val Val Cys Arg Ser Gly Ala Leu Thr Val Ser Glu Ile Ala Ala Ala
260 265 270

gga cta ccg gcg ttg ttt gtg ccg ttt caa cat aaa gac cgc cag caa 864
Gly Leu Pro Ala Leu Phe Val Pro Phe Gln His Lys Asp Arg Gln Gln
275 280 285

tac tgg aat gcg cta ccg ctg gaa aaa gcg ggc gca gcc aaa att atc 912
Tyr Trp Asn Ala Leu Pro Leu Glu Lys Ala Gly Ala Ala Lys Ile Ile
290 295 300

gag cag cca cag ctt agc gtg gat gct gtc gcc aac acc ctg gcc ggg 960
Glu Gln Pro Gln Leu Ser Val Asp Ala Val Ala Asn Thr Leu Ala Gly
305 310 315 320

tgg tcg cga gaa acc tta tta acc atg gca gaa cgc gcc cgc gct gca 1008
Trp Ser Arg Glu Thr Leu Leu Thr Met Ala Glu Arg Ala Arg Ala Ala
325 330 335

tcc att ccg gat gcc acc gag cga gtg gca aat gaa gtg agc cgg gtt 1056
Ser Ile Pro Asp Ala Thr Glu Arg Val Ala Asn Glu Val Ser Arg Val
340 345 350

gcc cgg gcg taa 1068
Ala Arg Ala *
355

<210> 262
<211> 1476
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1476)

<400> 262
atg aat aca caa caa ttg gca aaa ctg cgt tcc atc gtg ccc gaa atg 48
Met Asn Thr Gln Gln Leu Ala Lys Leu Arg Ser Ile Val Pro Glu Met
1 5 10 15

cgt cgc gtt cgg cac ata cat ttt gtc ggc att ggt ggt gcc ggt atg 96
Arg Arg Val Arg His Ile His Phe Val Gly Ile Gly Gly Ala Gly Met
20 25 30

ggc ggt att gcc gaa gtt ctg gcc aat gaa ggt tat cag atc agt ggt 144
Gly Gly Ile Ala Glu Val Leu Ala Asn Glu Gly Tyr Gln Ile Ser Gly
35 40 45

tcc gat tta gcg cca aat ccg gtc acg cag cag tta atg aat ctg ggt 192
Ser Asp Leu Ala Pro Asn Pro Val Thr Gln Gln Leu Met Asn Leu Gly
50 55 60

```
gcg acg att tat ttc aac cat cgc ccg gaa aac gta cgt gat gcc agc    240
Ala Thr Ile Tyr Phe Asn His Arg Pro Glu Asn Val Arg Asp Ala Ser
 65                 70                  75                  80

gtg gtc gtt gtt tcc agc gcg att tct gcc gat aac ccg gaa att gtc    288
Val Val Val Val Ser Ser Ala Ile Ser Ala Asp Asn Pro Glu Ile Val
                 85                  90                  95

gcc gct cat gaa gcg cgt att ccg gtg atc cgt cgt gcc gaa atg ctg    336
Ala Ala His Glu Ala Arg Ile Pro Val Ile Arg Arg Ala Glu Met Leu
            100                 105                 110

gct gag tta atg cgt ttt cgt cat ggc atc gcc att gcc gga acg cac    384
Ala Glu Leu Met Arg Phe Arg His Gly Ile Ala Ile Ala Gly Thr His
        115                 120                 125

ggc aaa acg aca acc acc gcg atg gtt tcc agc atc tac gca gaa gcg    432
Gly Lys Thr Thr Thr Thr Ala Met Val Ser Ser Ile Tyr Ala Glu Ala
    130                 135                 140

ggg ctc gac cca acc ttc gtt aac ggc ggg ctg gta aaa gcg gcg ggg    480
Gly Leu Asp Pro Thr Phe Val Asn Gly Gly Leu Val Lys Ala Ala Gly
145                 150                 155                 160

gtt cat gcg cgt ttg ggg cat ggt cgg tac ctg att gcc gaa gca gat    528
Val His Ala Arg Leu Gly His Gly Arg Tyr Leu Ile Ala Glu Ala Asp
                165                 170                 175

gag agt gat gca tcg ttc ctg cat ctg caa ccg atg gtg gcg att gtc    576
Glu Ser Asp Ala Ser Phe Leu His Leu Gln Pro Met Val Ala Ile Val
            180                 185                 190

acc aat atc gaa gcc gac cac atg gat acc tac cag ggc gac ttt gag    624
Thr Asn Ile Glu Ala Asp His Met Asp Thr Tyr Gln Gly Asp Phe Glu
        195                 200                 205

aat tta aaa cag act ttt att aat ttt ctg cac aac ctg ccg ttt tac    672
Asn Leu Lys Gln Thr Phe Ile Asn Phe Leu His Asn Leu Pro Phe Tyr
    210                 215                 220

ggt cgt gcg gtg atg tgt gtt gat gat ccg gtg atc cgc gaa ttg tta    720
Gly Arg Ala Val Met Cys Val Asp Asp Pro Val Ile Arg Glu Leu Leu
225                 230                 235                 240

ccg cga gtg ggg cgt cag acc acg act tac ggc ttc agc gaa gat gcc    768
Pro Arg Val Gly Arg Gln Thr Thr Thr Tyr Gly Phe Ser Glu Asp Ala
                245                 250                 255

gac gtg cgt gta gaa gat tat cag cag att ggc ccg cag ggg cac ttt    816
Asp Val Arg Val Glu Asp Tyr Gln Gln Ile Gly Pro Gln Gly His Phe
            260                 265                 270

acg ctg ctg cgc cag gac aaa gag ccg atg cgc gtc acc ctg aat gcg    864
Thr Leu Leu Arg Gln Asp Lys Glu Pro Met Arg Val Thr Leu Asn Ala
        275                 280                 285

cca ggt cgt cat aac gcg ctg aac gcc gca gct gcg gtt gcg gtt gct    912
Pro Gly Arg His Asn Ala Leu Asn Ala Ala Ala Ala Val Ala Val Ala
    290                 295                 300
```

acg gaa gag ggc att gac gac gag gct att ttg cgg gcg ctt gaa agc 960
Thr Glu Glu Gly Ile Asp Asp Glu Ala Ile Leu Arg Ala Leu Glu Ser
305 310 315 320

ttc cag ggg act ggt cgc cgt ttt gat ttc ctc ggt gaa ttc ccg ctg 1008
Phe Gln Gly Thr Gly Arg Arg Phe Asp Phe Leu Gly Glu Phe Pro Leu
325 330 335

gag cca gtg aat ggt aaa agc ggt acg gca atg ctg gtc gat gac tac 1056
Glu Pro Val Asn Gly Lys Ser Gly Thr Ala Met Leu Val Asp Asp Tyr
340 345 350

ggc cac cac ccg acg gaa gtg gac gcc acc att aaa gcg gcg cgc gca 1104
Gly His His Pro Thr Glu Val Asp Ala Thr Ile Lys Ala Ala Arg Ala
355 360 365

ggc tgg ccg gat aaa aac ctg gta atg ctg ttt cag ccg cac cgt ttt 1152
Gly Trp Pro Asp Lys Asn Leu Val Met Leu Phe Gln Pro His Arg Phe
370 375 380

acc cgt acg cgc gac ctg tat gat gat ttc gcc aat gtg ctg acg cag 1200
Thr Arg Thr Arg Asp Leu Tyr Asp Asp Phe Ala Asn Val Leu Thr Gln
385 390 395 400

gtt gat acc ctg ttg atg ctg gaa gtg tat ccg gct ggc gaa gcg cca 1248
Val Asp Thr Leu Leu Met Leu Glu Val Tyr Pro Ala Gly Glu Ala Pro
405 410 415

att ccg gga gcg gac agc cgt tcg ctg tgt cgc aca att cgt gga cgt 1296
Ile Pro Gly Ala Asp Ser Arg Ser Leu Cys Arg Thr Ile Arg Gly Arg
420 425 430

ggg aaa att gat ccc att ctg gtg ccg gat ccg gcg cgg gta gcc gag 1344
Gly Lys Ile Asp Pro Ile Leu Val Pro Asp Pro Ala Arg Val Ala Glu
435 440 445

atg ctg gca ccg gta tta acc ggt aac gac ctg att ctc gtt cag ggg 1392
Met Leu Ala Pro Val Leu Thr Gly Asn Asp Leu Ile Leu Val Gln Gly
450 455 460

gct ggt aat att gga aaa att gcc cgt tct tta gct gaa atc aaa ctg 1440
Ala Gly Asn Ile Gly Lys Ile Ala Arg Ser Leu Ala Glu Ile Lys Leu
465 470 475 480

aag ccg caa act ccg gag gaa gaa caa cat gac tga 1476
Lys Pro Gln Thr Pro Glu Glu Glu Gln His Asp *
485 490

<210> 263
<211> 921
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(921)

<400> 263
atg act gat aaa atc gcg gtc ctg ttg ggt ggg acc tcc gct gag cgg 48
Met Thr Asp Lys Ile Ala Val Leu Leu Gly Gly Thr Ser Ala Glu Arg

```
1               5                   10                  15

gaa gtt tct ctg aat tct ggc gca gcg gtg tta gcc gga ctg cgt gaa    96
Glu Val Ser Leu Asn Ser Gly Ala Ala Val Leu Ala Gly Leu Arg Glu
            20                  25                  30

ggc ggt att gac gcg tat cct gtc gac ccg aaa gaa gtc gac gtg acg   144
Gly Gly Ile Asp Ala Tyr Pro Val Asp Pro Lys Glu Val Asp Val Thr
        35                  40                  45

caa ctg aag tcg atg ggc ttt cag aaa gtg ttt atc gcg cta cac ggt   192
Gln Leu Lys Ser Met Gly Phe Gln Lys Val Phe Ile Ala Leu His Gly
    50                  55                  60

cgc ggc ggt gaa gat ggt acg ctg cag ggg atg ctc gag ctg atg ggc   240
Arg Gly Gly Glu Asp Gly Thr Leu Gln Gly Met Leu Glu Leu Met Gly
65                  70                  75                  80

ttg cct tat acc gga agc gga gtg atg gca tct gcg ctt tca atg gat   288
Leu Pro Tyr Thr Gly Ser Gly Val Met Ala Ser Ala Leu Ser Met Asp
                85                  90                  95

aaa cta cgc agc aaa ctt cta tgg caa ggt gcc ggt tta ccg gtc gcg   336
Lys Leu Arg Ser Lys Leu Leu Trp Gln Gly Ala Gly Leu Pro Val Ala
            100                 105                 110

ccg tgg gta gcg tta acc cgc gca gag ttt gaa aaa ggc ctg agc gat   384
Pro Trp Val Ala Leu Thr Arg Ala Glu Phe Glu Lys Gly Leu Ser Asp
        115                 120                 125

aag cag tta gca gaa att tct gct ctg ggt ttg ccg gtt atc gtt aag   432
Lys Gln Leu Ala Glu Ile Ser Ala Leu Gly Leu Pro Val Ile Val Lys
    130                 135                 140

ccg agc cgc gaa ggt tcc agt gtg gga atg tca aaa gta gta gca gaa   480
Pro Ser Arg Glu Gly Ser Ser Val Gly Met Ser Lys Val Val Ala Glu
145                 150                 155                 160

aat gct cta caa gat gca tta aga ttg gca ttt cag cac gat gaa gaa   528
Asn Ala Leu Gln Asp Ala Leu Arg Leu Ala Phe Gln His Asp Glu Glu
                165                 170                 175

gta ttg att gaa aaa tgg cta agt ggg ccg gag ttc acg gtt gcg ata   576
Val Leu Ile Glu Lys Trp Leu Ser Gly Pro Glu Phe Thr Val Ala Ile
            180                 185                 190

ctc ggt gaa gaa att tta ccg tca ata cgt att caa ccg tcc gga acc   624
Leu Gly Glu Glu Ile Leu Pro Ser Ile Arg Ile Gln Pro Ser Gly Thr
        195                 200                 205

ttc tat gat tat gag gcg aag tat ctc tct gat gag aca cag tat ttc   672
Phe Tyr Asp Tyr Glu Ala Lys Tyr Leu Ser Asp Glu Thr Gln Tyr Phe
    210                 215                 220

tgc ccc gca ggt ctg gaa gcg tca caa gag gcc aat ttg cag gca tta   720
Cys Pro Ala Gly Leu Glu Ala Ser Gln Glu Ala Asn Leu Gln Ala Leu
225                 230                 235                 240

gtg ctg aaa gca tgg acg acg tta ggt tgc aaa gga tgg gga cgt att   768
Val Leu Lys Ala Trp Thr Thr Leu Gly Cys Lys Gly Trp Gly Arg Ile
                245                 250                 255
```

gac gtt atg ctg gac agc gat gga cag ttt tat ctg ctg gaa gcc aat 816
Asp Val Met Leu Asp Ser Asp Gly Gln Phe Tyr Leu Leu Glu Ala Asn
260 265 270

acc tca ccg ggt atg acc agc cac agc ctg gtg ccg atg gcg gca cgt 864
Thr Ser Pro Gly Met Thr Ser His Ser Leu Val Pro Met Ala Ala Arg
275 280 285

cag gca ggt atg agc ttc tcg cag ttg gta gta cga att ctg gaa ctg 912
Gln Ala Gly Met Ser Phe Ser Gln Leu Val Val Arg Ile Leu Glu Leu
290 295 300

gcg gac taa 921
Ala Asp *
305

<210> 264
<211> 606
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(606)

<400> 264
atg gcg ctg cat gac gaa aac gtc gtc tgg cat agc cat ccg gtc act 48
Met Ala Leu His Asp Glu Asn Val Val Trp His Ser His Pro Val Thr
1 5 10 15

gtg caa caa cgc gag cta cac cac ggt cat cgt ggt gta gtg ctg tgg 96
Val Gln Gln Arg Glu Leu His His Gly His Arg Gly Val Val Leu Trp
20 25 30

ttt acc ggc ctc tcc ggg tcc ggt aaa tca acg gtc gcc ggg gcg ctg 144
Phe Thr Gly Leu Ser Gly Ser Gly Lys Ser Thr Val Ala Gly Ala Leu
35 40 45

gag gag gcg tta cat aaa ctc ggc gtc agt acg tat ctg ctg gat ggc 192
Glu Glu Ala Leu His Lys Leu Gly Val Ser Thr Tyr Leu Leu Asp Gly
50 55 60

gac aat gtt cgc cac gga tta tgc agc gat ctc ggt ttt agc gat gcc 240
Asp Asn Val Arg His Gly Leu Cys Ser Asp Leu Gly Phe Ser Asp Ala
65 70 75 80

gat cgt aaa gag aat atc cgt cgc gtc ggt gaa gtg gcg aat ttg atg 288
Asp Arg Lys Glu Asn Ile Arg Arg Val Gly Glu Val Ala Asn Leu Met
85 90 95

gtt gaa gcc gga ctg gtg gtg ctg acc gca ttt atc tcg cca cac cgc 336
Val Glu Ala Gly Leu Val Val Leu Thr Ala Phe Ile Ser Pro His Arg
100 105 110

gcc gaa cgc cag atg gtt cgc gaa cgc gta gga gaa ggg cgc ttt atc 384
Ala Glu Arg Gln Met Val Arg Glu Arg Val Gly Glu Gly Arg Phe Ile
115 120 125

gaa gtg ttt gtc gat acg ccg ctg gcg att tgc gaa gcc cgc gat ccc 432

```
Glu Val Phe Val Asp Thr Pro Leu Ala Ile Cys Glu Ala Arg Asp Pro
    130                 135                 140

aaa ggc tta tat aag aaa gcg cgt gcc ggt gaa ctg cgc aac ttt acg    480
Lys Gly Leu Tyr Lys Lys Ala Arg Ala Gly Glu Leu Arg Asn Phe Thr
145                 150                 155                 160

gga ata gat tcc gtt tac gaa gcg cct gaa tcg gca gaa att cat ctc    528
Gly Ile Asp Ser Val Tyr Glu Ala Pro Glu Ser Ala Glu Ile His Leu
                165                 170                 175

aat ggt gaa caa tta gta aca aat ttg gta cag caa tta tta gat ctg    576
Asn Gly Glu Gln Leu Val Thr Asn Leu Val Gln Gln Leu Leu Asp Leu
            180                 185                 190

ttg aga cag aac gat att atc aga tcc tga                            606
Leu Arg Gln Asn Asp Ile Ile Arg Ser  *
        195                 200


<210> 265
<211> 1428
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1428)

<400> 265
atg aac acc gca ctt gca caa caa atc gcc aat gaa ggc ggc gtc gaa    48
Met Asn Thr Ala Leu Ala Gln Gln Ile Ala Asn Glu Gly Gly Val Glu
 1               5                  10                  15

gcc tgg atg att gcg caa caa cat aaa agc ctg ctg cgt ttt ctg acc    96
Ala Trp Met Ile Ala Gln Gln His Lys Ser Leu Leu Arg Phe Leu Thr
            20                  25                  30

tgt ggt agc gtc gat gac ggc aaa agt act ctg att ggt cgt ctg ctg    144
Cys Gly Ser Val Asp Asp Gly Lys Ser Thr Leu Ile Gly Arg Leu Leu
        35                  40                  45

cac gat acc cgc caa atc tac gaa gat cag ctc tca tcg ctg cat aac    192
His Asp Thr Arg Gln Ile Tyr Glu Asp Gln Leu Ser Ser Leu His Asn
    50                  55                  60

gac agt aag cgt cac ggc acc cag ggc gaa aag ctg gat ctg gct ctg    240
Asp Ser Lys Arg His Gly Thr Gln Gly Glu Lys Leu Asp Leu Ala Leu
65                  70                  75                  80

ctg gtg gac ggc ctg caa gct gag cgc gaa cag ggc atc acc att gac    288
Leu Val Asp Gly Leu Gln Ala Glu Arg Glu Gln Gly Ile Thr Ile Asp
                85                  90                  95

gtg gcc tac cgc tat ttc tct acc gag aag cgt aaa ttt att atc gcc    336
Val Ala Tyr Arg Tyr Phe Ser Thr Glu Lys Arg Lys Phe Ile Ile Ala
            100                 105                 110

gac acc cca ggg cac gag cag tac acc cgc aat atg gcg act ggc gca    384
Asp Thr Pro Gly His Glu Gln Tyr Thr Arg Asn Met Ala Thr Gly Ala
        115                 120                 125
```

tcg aca tgt gaa ctg gcg atc tta ctg atc gat gcc cgt aaa ggc gtg 432
Ser Thr Cys Glu Leu Ala Ile Leu Leu Ile Asp Ala Arg Lys Gly Val
130 135 140

ctc gat caa acc cgt cgt cac agt ttt atc tcc aca ctg ttg ggg atc 480
Leu Asp Gln Thr Arg Arg His Ser Phe Ile Ser Thr Leu Leu Gly Ile
145 150 155 160

aaa cat ctg gtc gtg gcg atc aac aaa atg gat ctg gtg gat tac agt 528
Lys His Leu Val Val Ala Ile Asn Lys Met Asp Leu Val Asp Tyr Ser
165 170 175

gaa gag acg ttc acc cgt att cgt gaa gat tat ttg acc ttt gcc ggg 576
Glu Glu Thr Phe Thr Arg Ile Arg Glu Asp Tyr Leu Thr Phe Ala Gly
180 185 190

cag ctg ccg ggt aat ctg gat atc cgc ttt gtg ccg ctc tct gca ctg 624
Gln Leu Pro Gly Asn Leu Asp Ile Arg Phe Val Pro Leu Ser Ala Leu
195 200 205

gaa ggc gac aac gtg gca tcg caa agt gaa agt atg ccg tgg tac agc 672
Glu Gly Asp Asn Val Ala Ser Gln Ser Glu Ser Met Pro Trp Tyr Ser
210 215 220

ggt ccg aca ctg ctc gaa gtg ctg gaa acc gtg gag atc cag cga gtg 720
Gly Pro Thr Leu Leu Glu Val Leu Glu Thr Val Glu Ile Gln Arg Val
225 230 235 240

gtg gat gct cag cca atg cgc ttc ccg gtg cag tac gtt aat cgc ccg 768
Val Asp Ala Gln Pro Met Arg Phe Pro Val Gln Tyr Val Asn Arg Pro
245 250 255

aat ctc gat ttt cgt ggt tac gcc gga acg ctg gca tcc ggt cgc gtg 816
Asn Leu Asp Phe Arg Gly Tyr Ala Gly Thr Leu Ala Ser Gly Arg Val
260 265 270

gaa gtc ggg caa cgt gtc aaa gtg ctg ccc tct ggt gtg gaa tca aac 864
Glu Val Gly Gln Arg Val Lys Val Leu Pro Ser Gly Val Glu Ser Asn
275 280 285

gtc gcg cgg atc gtg act ttt gat ggt gat cgc gaa gaa gcc ttt gcc 912
Val Ala Arg Ile Val Thr Phe Asp Gly Asp Arg Glu Glu Ala Phe Ala
290 295 300

gga gaa gcg atc acc ctg gtg ctg acg gat gag atc gac atc agc cgt 960
Gly Glu Ala Ile Thr Leu Val Leu Thr Asp Glu Ile Asp Ile Ser Arg
305 310 315 320

ggc gat ctg ctg ctg gcg gca gac gaa gcg tta ccg gcg gtg cag agc 1008
Gly Asp Leu Leu Leu Ala Ala Asp Glu Ala Leu Pro Ala Val Gln Ser
325 330 335

gcg tcg gtg gat gtg gta tgg atg gcg gaa cag ccg ctt tct cca ggg 1056
Ala Ser Val Asp Val Val Trp Met Ala Glu Gln Pro Leu Ser Pro Gly
340 345 350

cag agt tac gac atc aaa att gcc ggt aag aag acg cgc gcg cgt gtt 1104
Gln Ser Tyr Asp Ile Lys Ile Ala Gly Lys Lys Thr Arg Ala Arg Val
355 360 365

```
gat ggc att cgc tat cag gtt gat att aat aac ctt acc cag cgt gaa    1152
Asp Gly Ile Arg Tyr Gln Val Asp Ile Asn Asn Leu Thr Gln Arg Glu
    370                 375                 380

gtt gaa aac ctg cca ctg aat ggg atc ggc ctc gtg gat ctc act ttt    1200
Val Glu Asn Leu Pro Leu Asn Gly Ile Gly Leu Val Asp Leu Thr Phe
385                 390                 395                 400

gac gag ccg ctg gtg tta gat cgt tat caa caa aat ccg gtg acg ggt    1248
Asp Glu Pro Leu Val Leu Asp Arg Tyr Gln Gln Asn Pro Val Thr Gly
                405                 410                 415

ggg ctg att ttt atc gat cgc ctg agc aat gtg acc gtg ggt gcc ggt    1296
Gly Leu Ile Phe Ile Asp Arg Leu Ser Asn Val Thr Val Gly Ala Gly
            420                 425                 430

atg gtg cac gag cca gtt agc cag gca act gct gcg cca tct gaa ttc    1344
Met Val His Glu Pro Val Ser Gln Ala Thr Ala Ala Pro Ser Glu Phe
        435                 440                 445

agt gca ttc gaa ctg gaa ttg aat gct ctg gtt cgt cgc cac ttt ccg    1392
Ser Ala Phe Glu Leu Glu Leu Asn Ala Leu Val Arg Arg His Phe Pro
    450                 455                 460

cac tgg ggc gcg cgc gat ttg ctg ggg gat aaa taa                    1428
His Trp Gly Ala Arg Asp Leu Leu Gly Asp Lys  *
465                 470                 475


<210> 266
<211> 384
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(384)

<400> 266
atg cgc cat cgt aag agt ggt cgt caa ctg aac cgc aac agc agc cat    48
Met Arg His Arg Lys Ser Gly Arg Gln Leu Asn Arg Asn Ser Ser His
 1               5                  10                  15

cgc cag gct atg ttc cgc aat atg gca ggt tca ctg gtt cgt cat gaa    96
Arg Gln Ala Met Phe Arg Asn Met Ala Gly Ser Leu Val Arg His Glu
            20                  25                  30

atc atc aag acg act ctg cct aaa gcg aaa gag ctg cgc cgc gta gtt    144
Ile Ile Lys Thr Thr Leu Pro Lys Ala Lys Glu Leu Arg Arg Val Val
        35                  40                  45

gag ccg ctg att act ctt gcc aag act gat agc gtt gct aat cgt cgt    192
Glu Pro Leu Ile Thr Leu Ala Lys Thr Asp Ser Val Ala Asn Arg Arg
    50                  55                  60

ctg gca ttc gcc cgt act cgt gat aac gag atc gtg gca aaa ctg ttt    240
Leu Ala Phe Ala Arg Thr Arg Asp Asn Glu Ile Val Ala Lys Leu Phe
 65                 70                  75                  80

aac gaa ctg ggc ccg cgt ttc gcg agc cgt gcc ggt ggt tac act cgt    288
Asn Glu Leu Gly Pro Arg Phe Ala Ser Arg Ala Gly Gly Tyr Thr Arg
```

```
                    85                  90                  95

att ctg aag tgt ggc ttc cgt gca ggc gac aac gcg ccg atg gct tac    336
Ile Leu Lys Cys Gly Phe Arg Ala Gly Asp Asn Ala Pro Met Ala Tyr
                100                 105                 110

atc gag ctg gtt gat cgt tca gag aaa gca gaa gct gct gca gag taa    384
Ile Glu Leu Val Asp Arg Ser Glu Lys Ala Glu Ala Ala Ala Glu  *
            115                 120                 125



<210> 267
<211> 990
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(990)

<400> 267
atg cag ggt tct gtg aca gag ttt cta aaa ccg cgc ctg gtt gat atc     48
Met Gln Gly Ser Val Thr Glu Phe Leu Lys Pro Arg Leu Val Asp Ile
 1               5                  10                  15

gag caa gtg agt tcg acg cac gcc aag gtg acc ctt gag cct tta gag     96
Glu Gln Val Ser Ser Thr His Ala Lys Val Thr Leu Glu Pro Leu Glu
                20                  25                  30

cgt ggc ttt ggc cat act ctg ggt aac gca ctg cgc cgt att ctg ctc    144
Arg Gly Phe Gly His Thr Leu Gly Asn Ala Leu Arg Arg Ile Leu Leu
            35                  40                  45

tca tcg atg ccg ggt tgc gcg gtg acc gag gtt gag att gat ggt gta    192
Ser Ser Met Pro Gly Cys Ala Val Thr Glu Val Glu Ile Asp Gly Val
        50                  55                  60

cta cat gag tac agc acc aaa gaa ggc gtt cag gaa gat atc ctg gaa    240
Leu His Glu Tyr Ser Thr Lys Glu Gly Val Gln Glu Asp Ile Leu Glu
 65                  70                  75                  80

atc ctg ctc aac ctg aaa ggg ctg gcg gtg aga gtt cag ggc aaa gat    288
Ile Leu Leu Asn Leu Lys Gly Leu Ala Val Arg Val Gln Gly Lys Asp
                    85                  90                  95

gaa gtt att ctt acc ttg aat aaa tct ggc att ggc cct gtg act gca    336
Glu Val Ile Leu Thr Leu Asn Lys Ser Gly Ile Gly Pro Val Thr Ala
                100                 105                 110

gcc gat atc acc cac gac ggt gat gtc gaa atc gtc aag ccg cag cac    384
Ala Asp Ile Thr His Asp Gly Asp Val Glu Ile Val Lys Pro Gln His
            115                 120                 125

gtg atc tgc cac ctg acc gat gag aac gcg tct att agc atg cgt atc    432
Val Ile Cys His Leu Thr Asp Glu Asn Ala Ser Ile Ser Met Arg Ile
        130                 135                 140

aaa gtt cag cgc ggt cgt ggt tat gtg ccg gct tct acc cga att cat    480
Lys Val Gln Arg Gly Arg Gly Tyr Val Pro Ala Ser Thr Arg Ile His
145                 150                 155                 160
```

```
tcg gaa gaa gat gag cgc cca atc ggc cgt ctg ctg gtc gac gca tgc     528
Ser Glu Glu Asp Glu Arg Pro Ile Gly Arg Leu Leu Val Asp Ala Cys
                165                 170                 175

tac agc cct gtg gag cgt att gcc tac aat gtt gaa gca gcg cgt gta     576
Tyr Ser Pro Val Glu Arg Ile Ala Tyr Asn Val Glu Ala Ala Arg Val
            180                 185                 190

gaa cag cgt acc gac ctg gac aag ctg gtc atc gaa atg gaa acc aac     624
Glu Gln Arg Thr Asp Leu Asp Lys Leu Val Ile Glu Met Glu Thr Asn
        195                 200                 205

ggc aca atc gat cct gaa gag gcg att cgt cgt gcg gca acc att ctg     672
Gly Thr Ile Asp Pro Glu Glu Ala Ile Arg Arg Ala Ala Thr Ile Leu
    210                 215                 220

gct gaa caa ctg gaa gct ttc gtt gac tta cgt gat gta cgt cag cct     720
Ala Glu Gln Leu Glu Ala Phe Val Asp Leu Arg Asp Val Arg Gln Pro
225                 230                 235                 240

gaa gtg aaa gaa gag aaa cca gag ttc gat ccg atc ctg ctg cgc cct     768
Glu Val Lys Glu Glu Lys Pro Glu Phe Asp Pro Ile Leu Leu Arg Pro
                245                 250                 255

gtt gac gat ctg gaa ttg act gtc cgc tct gct aac tgc ctt aaa gca     816
Val Asp Asp Leu Glu Leu Thr Val Arg Ser Ala Asn Cys Leu Lys Ala
            260                 265                 270

gaa gct atc cac tat atc ggt gat ctg gta cag cgt acc gag gtt gag     864
Glu Ala Ile His Tyr Ile Gly Asp Leu Val Gln Arg Thr Glu Val Glu
        275                 280                 285

ctc ctt aaa acg cct aac ctt ggt aaa aaa tct ctt act gag att aaa     912
Leu Leu Lys Thr Pro Asn Leu Gly Lys Lys Ser Leu Thr Glu Ile Lys
    290                 295                 300

gac gtg ctg gct tcc cgt gga ctg tct ctg ggc atg cgc ctg gaa aac     960
Asp Val Leu Ala Ser Arg Gly Leu Ser Leu Gly Met Arg Leu Glu Asn
305                 310                 315                 320

tgg cca ccg gca agc atc gct gac gag taa                             990
Trp Pro Pro Ala Ser Ile Ala Asp Glu  *
                325


<210> 268
<211> 621
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(621)

<400> 268
atg gca aga tat ttg ggt cct aag ctc aag ctg agc cgt cgt gag ggc     48
Met Ala Arg Tyr Leu Gly Pro Lys Leu Lys Leu Ser Arg Arg Glu Gly
 1               5                  10                  15

acc gac tta ttc ctt aag tct ggc gtt cgc gcg atc gat acc aag tgt     96
```

```
Thr Asp Leu Phe Leu Lys Ser Gly Val Arg Ala Ile Asp Thr Lys Cys
            20                  25                  30

aaa att gaa caa gct cct ggc cag cac ggt gcg cgt aaa ccg cgt ctg     144
Lys Ile Glu Gln Ala Pro Gly Gln His Gly Ala Arg Lys Pro Arg Leu
        35                  40                  45

tct gac tat ggt gtg cag ttg cgt gaa aag caa aaa gtt cgc cgt atc     192
Ser Asp Tyr Gly Val Gln Leu Arg Glu Lys Gln Lys Val Arg Arg Ile
    50                  55                  60

tat ggt gtg ctg gag cgt cag ttc cgt aac tac tac aaa gaa gca gca     240
Tyr Gly Val Leu Glu Arg Gln Phe Arg Asn Tyr Tyr Lys Glu Ala Ala
65                  70                  75                  80

cgt ctg aaa ggc aac acc ggt gaa aac ctg ttg gct ctg ctg gaa ggt     288
Arg Leu Lys Gly Asn Thr Gly Glu Asn Leu Leu Ala Leu Leu Glu Gly
                85                  90                  95

cgt ctg gac aac gtt gta tac cgt atg ggc ttc ggt gcc act cgt gca     336
Arg Leu Asp Asn Val Val Tyr Arg Met Gly Phe Gly Ala Thr Arg Ala
            100                 105                 110

gaa gca cgt cag ctg gtt agc cat aaa gca att atg gta aac ggt cgt     384
Glu Ala Arg Gln Leu Val Ser His Lys Ala Ile Met Val Asn Gly Arg
        115                 120                 125

gtt gtt aac atc gct tct tat cag gtt agt ccg aat gac gtt gta agc     432
Val Val Asn Ile Ala Ser Tyr Gln Val Ser Pro Asn Asp Val Val Ser
    130                 135                 140

att cgt gag aaa gcg aag aag cag tct cgc gtg aaa gcc gct ctg gag     480
Ile Arg Glu Lys Ala Lys Lys Gln Ser Arg Val Lys Ala Ala Leu Glu
145                 150                 155                 160

ctg gct gag cag cgt gaa aag cca acc tgg ctg gaa gtt gat gct ggc     528
Leu Ala Glu Gln Arg Glu Lys Pro Thr Trp Leu Glu Val Asp Ala Gly
                165                 170                 175

aag atg gaa ggt acg ttt aag cgt aag ccg gag cgt tct gat ctg tct     576
Lys Met Glu Gly Thr Phe Lys Arg Lys Pro Glu Arg Ser Asp Leu Ser
            180                 185                 190

gcg gac att aac gaa cac ctg atc gtc gag ctt tac tcc aag taa         621
Ala Asp Ile Asn Glu His Leu Ile Val Glu Leu Tyr Ser Lys  *
        195                 200                 205
```

<210> 269
<211> 390
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(390)

<400> 269

```
atg gca aag gca cca att cgt gca cgt aaa cgt gta aga aaa caa gtc     48
Met Ala Lys Ala Pro Ile Arg Ala Arg Lys Arg Val Arg Lys Gln Val
```

```
 1               5                   10                  15

tct gac ggc gtg gct cat atc cat gct tct ttc aac aac acc atc gtg    96
Ser Asp Gly Val Ala His Ile His Ala Ser Phe Asn Asn Thr Ile Val
            20                  25                  30

act atc act gat cgt cag ggt aac gcg ttg ggt tgg gca aca gcc ggt    144
Thr Ile Thr Asp Arg Gln Gly Asn Ala Leu Gly Trp Ala Thr Ala Gly
        35                  40                  45

ggt tcc ggt ttc cgt ggt tct cgc aaa tcc act ccg ttt gca gct cag    192
Gly Ser Gly Phe Arg Gly Ser Arg Lys Ser Thr Pro Phe Ala Ala Gln
    50                  55                  60

gtt gca gca gag cgt tgc gct gac gcc gtg aaa gaa tac ggc atc aag    240
Val Ala Ala Glu Arg Cys Ala Asp Ala Val Lys Glu Tyr Gly Ile Lys
 65                  70                  75                  80

aat ctg gaa gtt atg gtt aaa ggt ccg ggt cca ggc cgc gaa tct act    288
Asn Leu Glu Val Met Val Lys Gly Pro Gly Pro Gly Arg Glu Ser Thr
                85                  90                  95

att cgt gct ctg aac gcc gca ggt ttc cgc atc act aac att act gat    336
Ile Arg Ala Leu Asn Ala Ala Gly Phe Arg Ile Thr Asn Ile Thr Asp
            100                 105                 110

gtg act ccg atc cct cat aac ggt tgt cgt ccg ccg aaa aaa cgt cgc    384
Val Thr Pro Ile Pro His Asn Gly Cys Arg Pro Pro Lys Lys Arg Arg
        115                 120                 125

gta taa                                                            390
Val  *
```

<210> 270
<211> 357
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(357)

<400> 270

```
gtg gcc cgt ata gca ggc att aac att cct gat cat aag cat gcc gta    48
Met Ala Arg Ile Ala Gly Ile Asn Ile Pro Asp His Lys His Ala Val
 1               5                   10                  15

atc gca tta act tcg att tat ggc gtc ggc aag acc cgt tct aaa gcc    96
Ile Ala Leu Thr Ser Ile Tyr Gly Val Gly Lys Thr Arg Ser Lys Ala
            20                  25                  30

atc ctg gct gca gcg ggt atc gct gaa gat gtt aag atc agt gag ctg    144
Ile Leu Ala Ala Ala Gly Ile Ala Glu Asp Val Lys Ile Ser Glu Leu
        35                  40                  45

tct gaa gga caa atc gac acg ctg cgt gac gaa gtt gcc aaa ttt gtc    192
Ser Glu Gly Gln Ile Asp Thr Leu Arg Asp Glu Val Ala Lys Phe Val
    50                  55                  60
```

```
gtt gaa ggt gat ctg cgc cgt gaa atc agc atg agc atc aag cgc ctg    240
Val Glu Gly Asp Leu Arg Arg Glu Ile Ser Met Ser Ile Lys Arg Leu
 65                  70                  75                  80

atg gat ctt ggt tgc tat cgc ggt ttg cgt cat cgt cgt ggt ctc ccg    288
Met Asp Leu Gly Cys Tyr Arg Gly Leu Arg His Arg Arg Gly Leu Pro
                 85                  90                  95

gtt cgc ggt cag cgt acc aag acc aac gca cgt acc cgt aag ggt ccg    336
Val Arg Gly Gln Arg Thr Lys Thr Asn Ala Arg Thr Arg Lys Gly Pro
            100                 105                 110

cgc aaa ccg atc aag aaa taa                                        357
Arg Lys Pro Ile Lys Lys  *
        115


&lt;210&gt; 271
&lt;211&gt; 1383
&lt;212&gt; DNA
&lt;213&gt; Escherichia coli

&lt;220&gt;
&lt;221&gt; CDS
&lt;222&gt; (1)...(1383)

&lt;400&gt; 271
atg acg caa tta acc atg aaa gac aaa att ggc tac ggg ctg gga gac    48
Met Thr Gln Leu Thr Met Lys Asp Lys Ile Gly Tyr Gly Leu Gly Asp
 1               5                  10                  15

acc gcc tgc ggc ttc gtc tgg cag gcc acg atg ttc ctg ctg gcc tat    96
Thr Ala Cys Gly Phe Val Trp Gln Ala Thr Met Phe Leu Leu Ala Tyr
             20                  25                  30

ttc tac acc gac gtc ttc ggc ctg tcg gcg ggg att atg ggc acg ctg    144
Phe Tyr Thr Asp Val Phe Gly Leu Ser Ala Gly Ile Met Gly Thr Leu
         35                  40                  45

ttt ttg gtc tcc cgc gtg ctc gac gcc gtc acc gac ccg ctg atg ggg    192
Phe Leu Val Ser Arg Val Leu Asp Ala Val Thr Asp Pro Leu Met Gly
     50                  55                  60

ctg ctg gta gac cgc acc cgc acg cgg cac ggc cag ttc cgc ccg ttc    240
Leu Leu Val Asp Arg Thr Arg Thr Arg His Gly Gln Phe Arg Pro Phe
 65                  70                  75                  80

ctg ctg tgg ggg gcc atc ccg ttc ggc atc gtc tgc gtg ctg acc ttc    288
Leu Leu Trp Gly Ala Ile Pro Phe Gly Ile Val Cys Val Leu Thr Phe
                 85                  90                  95

tac acg ccg gac ttc tcc gca cag ggc aag atc atc tac gcc tgc gtg    336
Tyr Thr Pro Asp Phe Ser Ala Gln Gly Lys Ile Ile Tyr Ala Cys Val
            100                 105                 110

acc tac att ctc ctg acc ctg gtc tac acc ttc gtt aac gtg ccg tac    384
Thr Tyr Ile Leu Leu Thr Leu Val Tyr Thr Phe Val Asn Val Pro Tyr
        115                 120                 125

tgc gcc atg ccg ggc gtc atc acc gcc gac ccg aaa gag cgt cac gcc    432
Cys Ala Met Pro Gly Val Ile Thr Ala Asp Pro Lys Glu Arg His Ala
```

```
    130                 135                 140

ctg cag tcc tgg cgc ttc ttc ctg gcg gcg gcg ggc tcg ctc gct atc    480
Leu Gln Ser Trp Arg Phe Phe Leu Ala Ala Ala Gly Ser Leu Ala Ile
145                 150                 155                 160

agc ggc atc gcg ctg ccg ctg gtg agc atc atc ggc aaa ggg gac gag    528
Ser Gly Ile Ala Leu Pro Leu Val Ser Ile Ile Gly Lys Gly Asp Glu
                165                 170                 175

cag gtg ggc tac ttc ggc gcc atg tgc gtg ctg ggg ctg agc ggc gtg    576
Gln Val Gly Tyr Phe Gly Ala Met Cys Val Leu Gly Leu Ser Gly Val
            180                 185                 190

gtg ctg ctc tac gtc tgc ttc ttc acg acc aaa gag cgc tac acc ttt    624
Val Leu Leu Tyr Val Cys Phe Phe Thr Thr Lys Glu Arg Tyr Thr Phe
        195                 200                 205

gag gtg cag ccg ggc tcg tcg gtg gcg aaa gac ctt aag ctg ctg ctg    672
Glu Val Gln Pro Gly Ser Ser Val Ala Lys Asp Leu Lys Leu Leu Leu
    210                 215                 220

ggc aac agc cag tgg cgc atc atg tgc gcg ttc aag atg atg gcg acc    720
Gly Asn Ser Gln Trp Arg Ile Met Cys Ala Phe Lys Met Met Ala Thr
225                 230                 235                 240

tgc tcc aac gtg gtg cgc ggc ggg gcg acg ctc tac ttc gtg aaa tac    768
Cys Ser Asn Val Val Arg Gly Gly Ala Thr Leu Tyr Phe Val Lys Tyr
                245                 250                 255

gtg atg gat cac ccg gag ttg gcg acc cag ttt tta ctt tac ggc agc    816
Val Met Asp His Pro Glu Leu Ala Thr Gln Phe Leu Leu Tyr Gly Ser
            260                 265                 270

ctc gcc acc atg ttc ggc tcg ctt tgc tcc tca cgc ctg ctg ggc cgc    864
Leu Ala Thr Met Phe Gly Ser Leu Cys Ser Ser Arg Leu Leu Gly Arg
        275                 280                 285

ttc gac cgc gtc acc gcc ttc aag tgg atc atc gtc gcc tac tcg ctg    912
Phe Asp Arg Val Thr Ala Phe Lys Trp Ile Ile Val Ala Tyr Ser Leu
    290                 295                 300

atc agc ctg ctg att ttc gtc acc ccg gcg gag cac atc gcg ctc att    960
Ile Ser Leu Leu Ile Phe Val Thr Pro Ala Glu His Ile Ala Leu Ile
305                 310                 315                 320

ttt gcc ctc aac atc ctg ttc ctg ttc gtc ttt aat acc acc acg ccg    1008
Phe Ala Leu Asn Ile Leu Phe Leu Phe Val Phe Asn Thr Thr Thr Pro
                325                 330                 335

ctg cag tgg ctg atg gct tct gac gtg gtg gac tac gag gag agc cgc    1056
Leu Gln Trp Leu Met Ala Ser Asp Val Val Asp Tyr Glu Glu Ser Arg
            340                 345                 350

agc ggt cgc cgc ctc gac ggg ctg gtg ttc tcc acc tac ctg ttc agc    1104
Ser Gly Arg Arg Leu Asp Gly Leu Val Phe Ser Thr Tyr Leu Phe Ser
        355                 360                 365

ctg aag att ggc ctg gcg att ggc ggg gcg gtg gtg ggc tgg atc ctg    1152
Leu Lys Ile Gly Leu Ala Ile Gly Gly Ala Val Val Gly Trp Ile Leu
    370                 375                 380
```

gcg tac gtc aac tat tcc gcc agc agc agc gtg cag ccg gtt gag gtg 1200
Ala Tyr Val Asn Tyr Ser Ala Ser Ser Ser Val Gln Pro Val Glu Val
385 390 395 400

ctc acc acc atc aaa att ctg ttc tgc gtg gtg ccg gtg gtg ctc tac 1248
Leu Thr Thr Ile Lys Ile Leu Phe Cys Val Val Pro Val Val Leu Tyr
405 410 415

gcg ggc atg ttc atc atg ctg tcg ctc tac aag ctc acc gat gcc cgc 1296
Ala Gly Met Phe Ile Met Leu Ser Leu Tyr Lys Leu Thr Asp Ala Arg
420 425 430

gtg gag gcc atc agc cgg cag ctg att aag cac cgc gcg gcg cag ggc 1344
Val Glu Ala Ile Ser Arg Gln Leu Ile Lys His Arg Ala Ala Gln Gly
435 440 445

gag gcc gtt ccc gac gcc gcg aca gcc gca tcc cat taa 1383
Glu Ala Val Pro Asp Ala Ala Thr Ala Ala Ser His *
450 455 460

<210> 272
<211> 1611
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1611)

<400> 272
atg gaa atc act aac ccg ata ctc acc ggc ttc aac ccg gac ccg tcc 48
Met Glu Ile Thr Asn Pro Ile Leu Thr Gly Phe Asn Pro Asp Pro Ser
1 5 10 15

ctg tgc cgc cag ggc gag gac tac tac atc gcc acc tcg acc ttc gag 96
Leu Cys Arg Gln Gly Glu Asp Tyr Tyr Ile Ala Thr Ser Thr Phe Glu
20 25 30

tgg ttc ccg ggc gtg cgc atc tac cac tcc cgt gac ctg aaa aac tgg 144
Trp Phe Pro Gly Val Arg Ile Tyr His Ser Arg Asp Leu Lys Asn Trp
35 40 45

tcg ctg gtc agc acc ccg ttg gac cgc gtg tcg atg ctg gac atg aag 192
Ser Leu Val Ser Thr Pro Leu Asp Arg Val Ser Met Leu Asp Met Lys
50 55 60

ggc aac ccg gac tcc ggc ggc atc tgg gcg ccg tgc ctg agc tac gcc 240
Gly Asn Pro Asp Ser Gly Gly Ile Trp Ala Pro Cys Leu Ser Tyr Ala
65 70 75 80

gac ggt aaa ttc tgg ctg ctc tac acc gac gtg aag att gtc gac tcg 288
Asp Gly Lys Phe Trp Leu Leu Tyr Thr Asp Val Lys Ile Val Asp Ser
85 90 95

ccg tgg aaa aac ggc cgc aac ttc ctc gtc acc gcg ccc tcc atc gag 336
Pro Trp Lys Asn Gly Arg Asn Phe Leu Val Thr Ala Pro Ser Ile Glu
100 105 110

ggg cca tgg agc gag cca atc ccg atg ggc aac ggc ggg ttt gac ccg 384

Gly Pro Trp Ser Glu Pro Ile Pro Met Gly Asn Gly Gly Phe Asp Pro
115 120 125

tcc ctg ttc cac gac gac gat ggc cgc aaa tac tat atc tac cgc ccg 432
Ser Leu Phe His Asp Asp Asp Gly Arg Lys Tyr Tyr Ile Tyr Arg Pro
130 135 140

tgg ggg ccg cgc cac cac agc aac ccg cac aac acc atc gtg tta cag 480
Trp Gly Pro Arg His His Ser Asn Pro His Asn Thr Ile Val Leu Gln
145 150 155 160

gcg ttt gac ccg cag acc ggc acg ctc tcg ccc gag cgc aaa acg ctg 528
Ala Phe Asp Pro Gln Thr Gly Thr Leu Ser Pro Glu Arg Lys Thr Leu
165 170 175

ttt acc ggc acg ccg ctc tgc tac acc gaa ggc gcg cac ctg tat cgc 576
Phe Thr Gly Thr Pro Leu Cys Tyr Thr Glu Gly Ala His Leu Tyr Arg
180 185 190

cac gcg gga tgg tac tac ctg atg gcc gcc gag ggc ggc acc agc tac 624
His Ala Gly Trp Tyr Tyr Leu Met Ala Ala Glu Gly Gly Thr Ser Tyr
195 200 205

gag cac gcc gtc gtg gtg ctg cgt tcc aaa aat atc gac ggg ccg tac 672
Glu His Ala Val Val Val Leu Arg Ser Lys Asn Ile Asp Gly Pro Tyr
210 215 220

gag ctg cac ccg gac gta acg atg atg acc agc tgg cac ctg ccg gag 720
Glu Leu His Pro Asp Val Thr Met Met Thr Ser Trp His Leu Pro Glu
225 230 235 240

aac ccg ctg cag aag agc ggc cac ggc tcg ctg ctg cag acg cat acg 768
Asn Pro Leu Gln Lys Ser Gly His Gly Ser Leu Leu Gln Thr His Thr
245 250 255

ggt gaa tgg tac atg gcc tac ctc acc agc cgc ccg ctg cgc ctg ccc 816
Gly Glu Trp Tyr Met Ala Tyr Leu Thr Ser Arg Pro Leu Arg Leu Pro
260 265 270

ggc gtg ccg ctg ctg gcc tcc ggc gga cgc ggc tac tgc ccg ctg ggg 864
Gly Val Pro Leu Leu Ala Ser Gly Gly Arg Gly Tyr Cys Pro Leu Gly
275 280 285

cgc gag acc ggc atc gcc cgc att gaa tgg cgc gac ggc tgg ccg tac 912
Arg Glu Thr Gly Ile Ala Arg Ile Glu Trp Arg Asp Gly Trp Pro Tyr
290 295 300

gtg gaa ggc ggc aag cac gcg cag ctg acc gtg aaa ggc ccg caa gta 960
Val Glu Gly Gly Lys His Ala Gln Leu Thr Val Lys Gly Pro Gln Val
305 310 315 320

gcc gag cag cct gca gcc gtt ccg ggc aac tgg cgg gac gat ttc gac 1008
Ala Glu Gln Pro Ala Ala Val Pro Gly Asn Trp Arg Asp Asp Phe Asp
325 330 335

gcc agt tcg ctt gac ccg gag ctg cag acc ctg cgc att ccg ttc gac 1056
Ala Ser Ser Leu Asp Pro Glu Leu Gln Thr Leu Arg Ile Pro Phe Asp
340 345 350

gac acc ctc ggc tcg ctc acc gcg cgc ccg ggc ttc tta cgg ctc tat 1104
Asp Thr Leu Gly Ser Leu Thr Ala Arg Pro Gly Phe Leu Arg Leu Tyr

355 360 365

ggc aac gac tcg ctc aat tcg acc ttc acc caa tcg acc gtg gcg cgc 1152
Gly Asn Asp Ser Leu Asn Ser Thr Phe Thr Gln Ser Thr Val Ala Arg
370 375 380

cgc tgg cag cac ttc gcc ttc cgg gca gaa acg cgg atg gag ttc tcg 1200
Arg Trp Gln His Phe Ala Phe Arg Ala Glu Thr Arg Met Glu Phe Ser
385 390 395 400

ccg gtg cac ttc cag cag agc gcg ggg ctg acc tgc tac tac aac agc 1248
Pro Val His Phe Gln Gln Ser Ala Gly Leu Thr Cys Tyr Tyr Asn Ser
405 410 415

aaa aac tgg agc tac tgc ttt gtg gac tac gag gag gga cag ggt aga 1296
Lys Asn Trp Ser Tyr Cys Phe Val Asp Tyr Glu Glu Gly Gln Gly Arg
420 425 430

acc atc aaa gtt atc cag ctc gac cac aac gtg ccg tcg tgg ccg ctg 1344
Thr Ile Lys Val Ile Gln Leu Asp His Asn Val Pro Ser Trp Pro Leu
435 440 445

cac gag cag ccc att ccg gtg ccg gaa cat gcg gag agc gtc tgg ctg 1392
His Glu Gln Pro Ile Pro Val Pro Glu His Ala Glu Ser Val Trp Leu
450 455 460

cgg gtg gac gtg gat acg ctg gtc tac cgc tac agc tac tcg ttt gat 1440
Arg Val Asp Val Asp Thr Leu Val Tyr Arg Tyr Ser Tyr Ser Phe Asp
465 470 475 480

ggc gag acg tgg cac acc gtg ccg gtg acg tat gag gcg tgg aag ctg 1488
Gly Glu Thr Trp His Thr Val Pro Val Thr Tyr Glu Ala Trp Lys Leu
485 490 495

tcg gac gac tac atc ggc ggg cgc ggc ttc ttc acc ggc gcg ttt gtg 1536
Ser Asp Asp Tyr Ile Gly Gly Arg Gly Phe Phe Thr Gly Ala Phe Val
500 505 510

ggc ctg cac tgc gag gac atc agc ggc gac ggc tgc tac gcg gac ttc 1584
Gly Leu His Cys Glu Asp Ile Ser Gly Asp Gly Cys Tyr Ala Asp Phe
515 520 525

gac tac ttc acc tac gag ccg gtc taa 1611
Asp Tyr Phe Thr Tyr Glu Pro Val *
530 535

<210> 273
<211> 978
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(978)

<400> 273
ttg cat atg aaa aaa ata atc ttt gct ttt att ata tta ttt gtg ttt 48
Met His Met Lys Lys Ile Ile Phe Ala Phe Ile Ile Leu Phe Val Phe
1 5 10 15

tta ctc cct atg att att ttt tac caa ccc tgg gtc aat gca tta ccg 96
Leu Leu Pro Met Ile Ile Phe Tyr Gln Pro Trp Val Asn Ala Leu Pro
20 25 30

tca acg ccc cga cac gca agt cct gag caa tta gaa aaa aca gtt cgt 144
Ser Thr Pro Arg His Ala Ser Pro Glu Gln Leu Glu Lys Thr Val Arg
35 40 45

tat ctt aca caa act gtg cat cca cgt agt gcc gac aat att gat aat 192
Tyr Leu Thr Gln Thr Val His Pro Arg Ser Ala Asp Asn Ile Asp Asn
50 55 60

ctg aat agg tcc gca gaa tat ata aaa gag gtc ttt gtc agt agc ggt 240
Leu Asn Arg Ser Ala Glu Tyr Ile Lys Glu Val Phe Val Ser Ser Gly
65 70 75 80

gcc aga gtt acc tcg cag gac gtc ccc att acg ggt ggc ccc tac aaa 288
Ala Arg Val Thr Ser Gln Asp Val Pro Ile Thr Gly Gly Pro Tyr Lys
85 90 95

aac att gtt gct gat tat ggt cct gcc gat gga ccg ctg att att att 336
Asn Ile Val Ala Asp Tyr Gly Pro Ala Asp Gly Pro Leu Ile Ile Ile
100 105 110

ggt gcg cat tat gac tct gcc agc agt tat gaa aac gat caa ttg acc 384
Gly Ala His Tyr Asp Ser Ala Ser Ser Tyr Glu Asn Asp Gln Leu Thr
115 120 125

tat acg ccg ggc gcg gat gat aac gcc agc ggt gtg gca gga tta ctc 432
Tyr Thr Pro Gly Ala Asp Asp Asn Ala Ser Gly Val Ala Gly Leu Leu
130 135 140

gaa ctg gca cgt ttg tta cat cag caa gta ccg aaa aca ggc gtg cag 480
Glu Leu Ala Arg Leu Leu His Gln Gln Val Pro Lys Thr Gly Val Gln
145 150 155 160

ctg gtc gcc tat gcg tcg gaa gaa ccg ccc ttc ttt cgt agc gat gaa 528
Leu Val Ala Tyr Ala Ser Glu Glu Pro Pro Phe Phe Arg Ser Asp Glu
165 170 175

atg ggg agc gcg gtg cat gca gct tcg ctt gag cgt cca gtg aaa tta 576
Met Gly Ser Ala Val His Ala Ala Ser Leu Glu Arg Pro Val Lys Leu
180 185 190

atg ata gca ctg gag atg att ggc tat tac gac tct gcg cct gga agc 624
Met Ile Ala Leu Glu Met Ile Gly Tyr Tyr Asp Ser Ala Pro Gly Ser
195 200 205

cag aat tac cct tac ccg gca atg tcc tgg ctt tat ccc gat cgg gga 672
Gln Asn Tyr Pro Tyr Pro Ala Met Ser Trp Leu Tyr Pro Asp Arg Gly
210 215 220

gac ttt att gcc gtg gtc ggc aga ata cag gat atc aac gcc gtt cgt 720
Asp Phe Ile Ala Val Val Gly Arg Ile Gln Asp Ile Asn Ala Val Arg
225 230 235 240

cag gta aaa gcg gca ttg ttg tca tct cag gat tta tct gtt tat tct 768
Gln Val Lys Ala Ala Leu Leu Ser Ser Gln Asp Leu Ser Val Tyr Ser
245 250 255

atg aat acc cca ggg ttt att ccc ggt att gat ttc tct gac cac ctg 816

```
Met Asn Thr Pro Gly Phe Ile Pro Gly Ile Asp Phe Ser Asp His Leu
            260                 265                 270

aat tat tgg caa cac gat att ccc gcc ata atg att act gac acc gct      864
Asn Tyr Trp Gln His Asp Ile Pro Ala Ile Met Ile Thr Asp Thr Ala
        275                 280                 285

ttt tat cgt aat aaa caa tac cac ttg ccc ggt gat acc gca gac aga      912
Phe Tyr Arg Asn Lys Gln Tyr His Leu Pro Gly Asp Thr Ala Asp Arg
    290                 295                 300

ttg aat tat cag aaa atg gct cag gta gtg gat ggt gtt ata act ttg      960
Leu Asn Tyr Gln Lys Met Ala Gln Val Val Asp Gly Val Ile Thr Leu
305                 310                 315                 320

tta tac aac agt aaa taa                                              978
Leu Tyr Asn Ser Lys  *
                325


<210> 274
<211> 963
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(963)

<400> 274
gtg atg atc aaa acg cgt ttt tct cgc tgg cta acg ttt ttt acg ttc      48
Met Met Ile Lys Thr Arg Phe Ser Arg Trp Leu Thr Phe Phe Thr Phe
 1               5                  10                  15

gcc gct gcc gtg gcg ctg gcg cta ccg gca aaa gcc aac acc tgg ccg      96
Ala Ala Ala Val Ala Leu Ala Leu Pro Ala Lys Ala Asn Thr Trp Pro
            20                  25                  30

ctg ccg cca gcg ggc agt cgt ctg gtt ggc gaa aac aaa ttt cat gtg      144
Leu Pro Pro Ala Gly Ser Arg Leu Val Gly Glu Asn Lys Phe His Val
        35                  40                  45

gtg gaa aat gac ggt ggt tct ctg gaa gcc atc gcc aaa aaa tac aac      192
Val Glu Asn Asp Gly Gly Ser Leu Glu Ala Ile Ala Lys Lys Tyr Asn
    50                  55                  60

gtc ggc ttt ctc gct ctg tta cag gct aac ccc ggc gtt gat cct tac      240
Val Gly Phe Leu Ala Leu Leu Gln Ala Asn Pro Gly Val Asp Pro Tyr
65                  70                  75                  80

gta ccg cgc gcg ggc agc gtg tta acg atc ccg ttg caa acc cta ctt      288
Val Pro Arg Ala Gly Ser Val Leu Thr Ile Pro Leu Gln Thr Leu Leu
                85                  90                  95

cca gat gcg ccg cgc gaa ggc att gtg atc aac att gcg gag ctg cgt      336
Pro Asp Ala Pro Arg Glu Gly Ile Val Ile Asn Ile Ala Glu Leu Arg
            100                 105                 110

ctc tat tac tac ccg ccg ggt aaa aat tcg gta acc gtg tat cca ata      384
Leu Tyr Tyr Tyr Pro Pro Gly Lys Asn Ser Val Thr Val Tyr Pro Ile
        115                 120                 125
```

```
ggt att ggt cag tta ggt ggt gac acg ctg aca ccg aca atg gtg acc     432
Gly Ile Gly Gln Leu Gly Gly Asp Thr Leu Thr Pro Thr Met Val Thr
    130                 135                 140

acc gtt tca gac aaa cgt gca aac cca acc tgg acg cca acg gca aac     480
Thr Val Ser Asp Lys Arg Ala Asn Pro Thr Trp Thr Pro Thr Ala Asn
145                 150                 155                 160

atc cgc gcc cgt tat aaa gca cag gga att gag ttg cct gcg gta gtg     528
Ile Arg Ala Arg Tyr Lys Ala Gln Gly Ile Glu Leu Pro Ala Val Val
                165                 170                 175

ccg gct gga ctg gat aac cca atg ggc cat cat gcg att cgt ctg gcg     576
Pro Ala Gly Leu Asp Asn Pro Met Gly His His Ala Ile Arg Leu Ala
            180                 185                 190

gcc tat ggc ggc gtt tat ttg ctt cat ggt acg aac gcc gat ttc ggc     624
Ala Tyr Gly Gly Val Tyr Leu Leu His Gly Thr Asn Ala Asp Phe Gly
        195                 200                 205

att ggc atg cgg gta agt tct ggc tgt att cgt ctg cgg gat gac gat     672
Ile Gly Met Arg Val Ser Ser Gly Cys Ile Arg Leu Arg Asp Asp Asp
    210                 215                 220

atc aaa aca ctc ttt agc cag gtc acc cca ggc acc aaa gtg aat atc     720
Ile Lys Thr Leu Phe Ser Gln Val Thr Pro Gly Thr Lys Val Asn Ile
225                 230                 235                 240

atc aac act ccg ata aaa gtc tct gcc gaa cca aac ggt gcg cgt ctg     768
Ile Asn Thr Pro Ile Lys Val Ser Ala Glu Pro Asn Gly Ala Arg Leu
                245                 250                 255

gtt gaa gta cat cag ccg ctg tca gag aag att gat gac gat ccg cag     816
Val Glu Val His Gln Pro Leu Ser Glu Lys Ile Asp Asp Asp Pro Gln
            260                 265                 270

ctg ctg cca att acg ctg aat agc gca atg caa tca ttt aaa gat gca     864
Leu Leu Pro Ile Thr Leu Asn Ser Ala Met Gln Ser Phe Lys Asp Ala
        275                 280                 285

gca caa act gac gct gaa gtg atg caa cat gtg atg gat gtc cgt tcc     912
Ala Gln Thr Asp Ala Glu Val Met Gln His Val Met Asp Val Arg Ser
    290                 295                 300

ggg atg ccg gtg gat gtc cgc cgt cat caa gtg agc cca caa acg ctg     960
Gly Met Pro Val Asp Val Arg Arg His Gln Val Ser Pro Gln Thr Leu
305                 310                 315                 320

taa                                                                 963
 *
```

<210> 275
<211> 1485
<212> DNA
<213> Escherichia coli

<220>
<221> CDS

<222> (1)...(1485)

<400> 275

```
gtg gtc gca atc cat ctt ttg ccg gtt agt tac aat tct gcg aca tcc     48
Met Val Ala Ile His Leu Leu Pro Val Ser Tyr Asn Ser Ala Thr Ser
 1               5                  10                  15

acc gtg aat atc agt gct aga atc ata ccc ctg ttg att att cac caa     96
Thr Val Asn Ile Ser Ala Arg Ile Ile Pro Leu Leu Ile Ile His Gln
            20                  25                  30

aga tat aaa att cct atg cca aaa gtt cag gcc gac ggc ctg cca ttg    144
Arg Tyr Lys Ile Pro Met Pro Lys Val Gln Ala Asp Gly Leu Pro Leu
        35                  40                  45

ccc cag cga tac ggt gcg ata tta acc att gtg att ggt att tcg atg    192
Pro Gln Arg Tyr Gly Ala Ile Leu Thr Ile Val Ile Gly Ile Ser Met
    50                  55                  60

gcc gtc ctt gac ggc gca atc gcc aac gtc gcc ctg cca aca atc gcc    240
Ala Val Leu Asp Gly Ala Ile Ala Asn Val Ala Leu Pro Thr Ile Ala
65                  70                  75                  80

acg gac ctt cat gcc acg cca gcc agt tcc atc tgg gta gtg aac gcc    288
Thr Asp Leu His Ala Thr Pro Ala Ser Ser Ile Trp Val Val Asn Ala
                85                  90                  95

tat caa atc gcc att gtc atc tcc ctg ctc tcg ttt tcg ttt ctg ggc    336
Tyr Gln Ile Ala Ile Val Ile Ser Leu Leu Ser Phe Ser Phe Leu Gly
            100                 105                 110

gat atg ttt ggc tat cga cgt att tat aaa tgc ggt ctg gtc gtt ttt    384
Asp Met Phe Gly Tyr Arg Arg Ile Tyr Lys Cys Gly Leu Val Val Phe
        115                 120                 125

ctg ttg tct tca ctg ttc tgc gcc ctt tct gat tcg ctg caa atg ctc    432
Leu Leu Ser Ser Leu Phe Cys Ala Leu Ser Asp Ser Leu Gln Met Leu
    130                 135                 140

acc ctt gcg cgt gtc ata caa ggt ttc ggc ggt gca gcg ttg atg agc    480
Thr Leu Ala Arg Val Ile Gln Gly Phe Gly Gly Ala Ala Leu Met Ser
145                 150                 155                 160

gtt aat acc gca ctt atc cgc ctg atc tat cca caa cgt ttt ctg ggt    528
Val Asn Thr Ala Leu Ile Arg Leu Ile Tyr Pro Gln Arg Phe Leu Gly
                165                 170                 175

aga ggg atg ggc ata aac tcg ttt att gtt gcc gtc tct tct gct gcc    576
Arg Gly Met Gly Ile Asn Ser Phe Ile Val Ala Val Ser Ser Ala Ala
            180                 185                 190

ggg ccg aca att gct gca gca atc ctc tcc atc gca tcc tgg aaa tgg    624
Gly Pro Thr Ile Ala Ala Ala Ile Leu Ser Ile Ala Ser Trp Lys Trp
        195                 200                 205

tta ttt tta atc aac gta ccg tta ggt att atc gcc ctg ctt ctg gcg    672
Leu Phe Leu Ile Asn Val Pro Leu Gly Ile Ile Ala Leu Leu Leu Ala
    210                 215                 220

atg cgt ttt ctg cca ccc aat ggt tct cgc gcc agt aaa ccc cgt ttc    720
Met Arg Phe Leu Pro Pro Asn Gly Ser Arg Ala Ser Lys Pro Arg Phe
```

225 230 235 240

gac ctg ccc agc gcc gtg atg aac gcg tta acc ttc ggc ctg ctt atc 768
Asp Leu Pro Ser Ala Val Met Asn Ala Leu Thr Phe Gly Leu Leu Ile
245 250 255

act gcg ttg agt ggt ttc gct cag ggg caa tcg ctg acg tta att gct 816
Thr Ala Leu Ser Gly Phe Ala Gln Gly Gln Ser Leu Thr Leu Ile Ala
260 265 270

gcg gaa ctg gtg gta atg gtt gtt gtt ggt att ttc ttt att cgc cgc 864
Ala Glu Leu Val Val Met Val Val Val Gly Ile Phe Phe Ile Arg Arg
275 280 285

cag ctt tct ctt ccc gta ccg ctg cta ccg gtg gat tta ctg cgt atc 912
Gln Leu Ser Leu Pro Val Pro Leu Leu Pro Val Asp Leu Leu Arg Ile
290 295 300

ccg ctg ttt tca ctt tct att tgc aca tct gtt tgc tct ttc tgc gca 960
Pro Leu Phe Ser Leu Ser Ile Cys Thr Ser Val Cys Ser Phe Cys Ala
305 310 315 320

caa atg ctg gca atg gtt tcc ctg ccc ttt tac ctg caa acc gtg ctc 1008
Gln Met Leu Ala Met Val Ser Leu Pro Phe Tyr Leu Gln Thr Val Leu
325 330 335

ggg cgt agt gaa gtc gaa aca ggt tta ctt ctg aca ccg tgg ccg tta 1056
Gly Arg Ser Glu Val Glu Thr Gly Leu Leu Leu Thr Pro Trp Pro Leu
340 345 350

gca acg atg gtg atg gct ccg ctg gca ggc tat ttg att gaa cgc gta 1104
Ala Thr Met Val Met Ala Pro Leu Ala Gly Tyr Leu Ile Glu Arg Val
355 360 365

cat gca gga ttg ctg ggg gct tta ggg ttg ttc atc atg gct gcg ggg 1152
His Ala Gly Leu Leu Gly Ala Leu Gly Leu Phe Ile Met Ala Ala Gly
370 375 380

ctt ttt tcc ctg gtt ctg ctg ccc gcg tca cct gcg gat atc aat att 1200
Leu Phe Ser Leu Val Leu Leu Pro Ala Ser Pro Ala Asp Ile Asn Ile
385 390 395 400

atc tgg ccg atg atc tta tgt ggt gct gga ttt ggc tta ttc cag tca 1248
Ile Trp Pro Met Ile Leu Cys Gly Ala Gly Phe Gly Leu Phe Gln Ser
405 410 415

ccc aat aac cac acc att att acc tcc gcg cct cgc gaa cgt agc ggt 1296
Pro Asn Asn His Thr Ile Ile Thr Ser Ala Pro Arg Glu Arg Ser Gly
420 425 430

gga gcc agt ggc atg tta gga acg gct cgt cta ctg ggt cag agt agc 1344
Gly Ala Ser Gly Met Leu Gly Thr Ala Arg Leu Leu Gly Gln Ser Ser
435 440 445

ggc gcg gcg ctg gtg gcg ctg atg cta aat cag ttt gga gat aat ggt 1392
Gly Ala Ala Leu Val Ala Leu Met Leu Asn Gln Phe Gly Asp Asn Gly
450 455 460

aca cac gtc tcg ctg atg gct gcg gct att ctg gca gtg att gct gcc 1440
Thr His Val Ser Leu Met Ala Ala Ala Ile Leu Ala Val Ile Ala Ala
465 470 475 480

tgt gtc agt ggt tta cgt atc act cag cca cga tcc agg gca taa 1485
Cys Val Ser Gly Leu Arg Ile Thr Gln Pro Arg Ser Arg Ala *
485 490

<210> 276
<211> 1434
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1434)

<400> 276
atg aaa gta acg ctg cca gag ttt gaa cgt gca gga gtg atg gtg gtt 48
Met Lys Val Thr Leu Pro Glu Phe Glu Arg Ala Gly Val Met Val Val
1 5 10 15

ggt gat gtg atg ctg gat cgt tac tgg tac ggc ccc acc agt cgt atc 96
Gly Asp Val Met Leu Asp Arg Tyr Trp Tyr Gly Pro Thr Ser Arg Ile
20 25 30

tcg ccg gaa gcg ccg gtg ccc gtg gtt aaa gtg aat acc atc gaa gaa 144
Ser Pro Glu Ala Pro Val Pro Val Val Lys Val Asn Thr Ile Glu Glu
35 40 45

cgt ccg ggc ggc gcg gct aac gtg gcg atg aat atc gct tct ctc ggt 192
Arg Pro Gly Gly Ala Ala Asn Val Ala Met Asn Ile Ala Ser Leu Gly
50 55 60

gct aat gca cgc ctg gtc ggg ttg acg ggc att gac gat gca gcg cgc 240
Ala Asn Ala Arg Leu Val Gly Leu Thr Gly Ile Asp Asp Ala Ala Arg
65 70 75 80

gcg ctg agt aaa tct ctg gcc gac gtc aac gtc aaa tgc gac ttc gtt 288
Ala Leu Ser Lys Ser Leu Ala Asp Val Asn Val Lys Cys Asp Phe Val
85 90 95

tct gta ccg acg cat ccg acc att acc aaa tta cgg gta ctt tcc cgc 336
Ser Val Pro Thr His Pro Thr Ile Thr Lys Leu Arg Val Leu Ser Arg
100 105 110

aac caa cag ctg atc cgt ctg gat ttt gaa gaa ggt ttc gaa ggt gtt 384
Asn Gln Gln Leu Ile Arg Leu Asp Phe Glu Glu Gly Phe Glu Gly Val
115 120 125

gat ccg cag ccg ctg cac gag cgg att aat cag gcg ctg agt tcg att 432
Asp Pro Gln Pro Leu His Glu Arg Ile Asn Gln Ala Leu Ser Ser Ile
130 135 140

ggc gcg ctg gtg ctt tct gac tac gcc aaa ggt gcg ctg gca agc gta 480
Gly Ala Leu Val Leu Ser Asp Tyr Ala Lys Gly Ala Leu Ala Ser Val
145 150 155 160

cag cag atg atc caa ctg gcg cgt aaa gcg ggt gtt ccg gtg ctg att 528
Gln Gln Met Ile Gln Leu Ala Arg Lys Ala Gly Val Pro Val Leu Ile
165 170 175

```
gat cca aaa ggt acc gat ttt gag cgc tac cgc ggc gct acg ctg tta    576
Asp Pro Lys Gly Thr Asp Phe Glu Arg Tyr Arg Gly Ala Thr Leu Leu
            180                 185                 190

acg ccg aat ctc tcg gaa ttt gaa gct gtt gtc ggt aaa tgt aag acc    624
Thr Pro Asn Leu Ser Glu Phe Glu Ala Val Val Gly Lys Cys Lys Thr
        195                 200                 205

gaa gaa gag att gtt gag cgc ggc atg aaa ctg att gcc gat tac gaa    672
Glu Glu Glu Ile Val Glu Arg Gly Met Lys Leu Ile Ala Asp Tyr Glu
    210                 215                 220

ctc tcg gct ctg tta gtg acc cgt tcc gaa cag ggt atg tcg ctg ctg    720
Leu Ser Ala Leu Leu Val Thr Arg Ser Glu Gln Gly Met Ser Leu Leu
225                 230                 235                 240

caa ccg ggt aaa gcg ccg ctg cat atg cca acc caa gcg cag gaa gtg    768
Gln Pro Gly Lys Ala Pro Leu His Met Pro Thr Gln Ala Gln Glu Val
                245                 250                 255

tat gac gtt acc ggt gcg ggc gac acg gtg att ggc gtc ctg gcg gca    816
Tyr Asp Val Thr Gly Ala Gly Asp Thr Val Ile Gly Val Leu Ala Ala
            260                 265                 270

acg ctg gca gcg ggt aat tcg ctg gaa gaa gcc tgc ttc ttt gcc aat    864
Thr Leu Ala Ala Gly Asn Ser Leu Glu Glu Ala Cys Phe Phe Ala Asn
        275                 280                 285

gcg gcg gct ggc gtg gtg gtc ggc aaa ctg gga acc tcc acg gtt tcg    912
Ala Ala Ala Gly Val Val Val Gly Lys Leu Gly Thr Ser Thr Val Ser
    290                 295                 300

ccg atc gag ctg gaa aat gct gta cgt gga cgt gca gat aca ggc ttt    960
Pro Ile Glu Leu Glu Asn Ala Val Arg Gly Arg Ala Asp Thr Gly Phe
305                 310                 315                 320

ggc gtg atg acc gaa gag gaa ctg aag ctg gcc gta gcg gca gcg cgt    1008
Gly Val Met Thr Glu Glu Glu Leu Lys Leu Ala Val Ala Ala Ala Arg
                325                 330                 335

aaa cgt ggt gaa aaa gtg gtg atg acc aac ggt gtc ttt gac atc ctg    1056
Lys Arg Gly Glu Lys Val Val Met Thr Asn Gly Val Phe Asp Ile Leu
            340                 345                 350

cac gcc ggg cac gtc tct tat ctg gca aat gcc cgc aag ctg ggt gac    1104
His Ala Gly His Val Ser Tyr Leu Ala Asn Ala Arg Lys Leu Gly Asp
        355                 360                 365

cgc ttg att gtt gcc gtc aac agc gat gcc tcc acc aaa cgg ctg aaa    1152
Arg Leu Ile Val Ala Val Asn Ser Asp Ala Ser Thr Lys Arg Leu Lys
    370                 375                 380

ggg gat tcc cgc ccg gta aac cca ctc gaa cag cgt atg att gtg ctg    1200
Gly Asp Ser Arg Pro Val Asn Pro Leu Glu Gln Arg Met Ile Val Leu
385                 390                 395                 400

ggc gca ctg gaa gcg gtc gac tgg gta gtg tcg ttt gaa gag gac acg    1248
Gly Ala Leu Glu Ala Val Asp Trp Val Val Ser Phe Glu Glu Asp Thr
                405                 410                 415

ccg cag cgc ttg atc gcc ggg atc ttg cca gat ctg ctg gtg aaa ggc    1296
```

Pro Gln Arg Leu Ile Ala Gly Ile Leu Pro Asp Leu Leu Val Lys Gly
420 425 430

ggc gac tat aaa cca gaa gag att gcc ggg agt aaa gaa gtc tgg gcc 1344
Gly Asp Tyr Lys Pro Glu Glu Ile Ala Gly Ser Lys Glu Val Trp Ala
435 440 445

aac ggt ggc gaa gtg ttg gtg ctc aac ttt gaa gac ggt tgc tcg acg 1392
Asn Gly Gly Glu Val Leu Val Leu Asn Phe Glu Asp Gly Cys Ser Thr
450 455 460

acc aac atc atc aag aag atc caa cag gat aaa aaa ggc taa 1434
Thr Asn Ile Ile Lys Lys Ile Gln Gln Asp Lys Lys Gly *
465 470 475

<210> 277
<211> 2841
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(2841)

<400> 277
atg aag ccg ctc tct tca ccg tta cag cag tac tgg cag acc gtt gtt 48
Met Lys Pro Leu Ser Ser Pro Leu Gln Gln Tyr Trp Gln Thr Val Val
1 5 10 15

gag cgg ctg cca gag cct tta gcc gag gaa tca ctt agc gca cag gcg 96
Glu Arg Leu Pro Glu Pro Leu Ala Glu Glu Ser Leu Ser Ala Gln Ala
20 25 30

aag tca gta ctt act ttt agt gat ttt gtg cag gac agc gtg att gcg 144
Lys Ser Val Leu Thr Phe Ser Asp Phe Val Gln Asp Ser Val Ile Ala
35 40 45

cat cca gag tgg ctg acg gaa ctg gaa agc caa ccg ccg cag gcc gac 192
His Pro Glu Trp Leu Thr Glu Leu Glu Ser Gln Pro Pro Gln Ala Asp
50 55 60

gaa tgg cag cat tac gcg gca tgg ttg cag gag gcg ctc tgt aat gtg 240
Glu Trp Gln His Tyr Ala Ala Trp Leu Gln Glu Ala Leu Cys Asn Val
65 70 75 80

agt gac gaa gcc ggg tta atg cgc gag ctg cgg cta ttc cgg cgg cgc 288
Ser Asp Glu Ala Gly Leu Met Arg Glu Leu Arg Leu Phe Arg Arg Arg
85 90 95

att atg gtg cgc atc gcc tgg gcg caa acg ctg gca ctg gtt act gaa 336
Ile Met Val Arg Ile Ala Trp Ala Gln Thr Leu Ala Leu Val Thr Glu
100 105 110

gag agc ata ttg cag cag ctc agc tat ctg gcg gag acg ctg att gtt 384
Glu Ser Ile Leu Gln Gln Leu Ser Tyr Leu Ala Glu Thr Leu Ile Val
115 120 125

gcg gcg cgt gac tgg ctg tat gac gcc tgc tgc cgc gag tgg gga acg 432
Ala Ala Arg Asp Trp Leu Tyr Asp Ala Cys Cys Arg Glu Trp Gly Thr

```
    130                 135                 140

ccg tgc aat gcg cag ggc gaa gcg caa ccg ctg ctg att tta ggc atg      480
Pro Cys Asn Ala Gln Gly Glu Ala Gln Pro Leu Leu Ile Leu Gly Met
145                 150                 155                 160

ggt aag ctg ggc ggt ggg gag ctg aat ttc tcc tct gat atc gat ctg      528
Gly Lys Leu Gly Gly Gly Glu Leu Asn Phe Ser Ser Asp Ile Asp Leu
                165                 170                 175

att ttt gcc tgg ccg gaa cat ggt tgt acg cag ggt gga cgc cgg gaa      576
Ile Phe Ala Trp Pro Glu His Gly Cys Thr Gln Gly Gly Arg Arg Glu
            180                 185                 190

ctg gat aac gcg cag ttt ttt acc cgc atg ggg cag cgg ctg att aaa      624
Leu Asp Asn Ala Gln Phe Phe Thr Arg Met Gly Gln Arg Leu Ile Lys
        195                 200                 205

gtg ctg gat caa cca acg cag gat ggc ttc gtc tat cgc gtg gat atg      672
Val Leu Asp Gln Pro Thr Gln Asp Gly Phe Val Tyr Arg Val Asp Met
    210                 215                 220

cgg ctg cgt ccg ttt ggc gaa agt ggc ccg ctg gtg ctg agc ttt gcc      720
Arg Leu Arg Pro Phe Gly Glu Ser Gly Pro Leu Val Leu Ser Phe Ala
225                 230                 235                 240

gcg ttg gaa gat tat tac cag gag cag ggg cgc gac tgg gag cgt tac      768
Ala Leu Glu Asp Tyr Tyr Gln Glu Gln Gly Arg Asp Trp Glu Arg Tyr
                245                 250                 255

gcg atg gtc aag gcg cgg att atg ggc gat agc gaa ggc gtc tat gct      816
Ala Met Val Lys Ala Arg Ile Met Gly Asp Ser Glu Gly Val Tyr Ala
            260                 265                 270

aac gag ttg cgt gcg atg ctg cgc ccg ttt gtt ttc cgt cgt tac atc      864
Asn Glu Leu Arg Ala Met Leu Arg Pro Phe Val Phe Arg Arg Tyr Ile
        275                 280                 285

gat ttc agc gtg att cag tcg ctg cgc aac atg aaa ggg atg att gcc      912
Asp Phe Ser Val Ile Gln Ser Leu Arg Asn Met Lys Gly Met Ile Ala
    290                 295                 300

cgt gaa gtg cgt cga cgt ggt ttg acc gac aat atc aaa ctc ggc gca      960
Arg Glu Val Arg Arg Arg Gly Leu Thr Asp Asn Ile Lys Leu Gly Ala
305                 310                 315                 320

ggc ggc att cgc gaa att gaa ttt atc gtt cag gtg ttc cag ctc att     1008
Gly Gly Ile Arg Glu Ile Glu Phe Ile Val Gln Val Phe Gln Leu Ile
                325                 330                 335

cgc ggc gga cgc gaa ccg tcg ctg caa tcg cgc tct tta ctg cca acg     1056
Arg Gly Gly Arg Glu Pro Ser Leu Gln Ser Arg Ser Leu Leu Pro Thr
            340                 345                 350

ctc agc gcc att gcc gag ctg cat ctg ctt tct gaa aac gat gct gaa     1104
Leu Ser Ala Ile Ala Glu Leu His Leu Leu Ser Glu Asn Asp Ala Glu
        355                 360                 365

caa ttg cga gtg gcg tat ctg ttc ctg cgg cgt ctg gaa aac ctg ctg     1152
Gln Leu Arg Val Ala Tyr Leu Phe Leu Arg Arg Leu Glu Asn Leu Leu
    370                 375                 380
```

caa agc att aac gac gaa caa acc cag acg ctt cct tct gat gag ctt 1200
Gln Ser Ile Asn Asp Glu Gln Thr Gln Thr Leu Pro Ser Asp Glu Leu
385 390 395 400

aat cgt gcg cgg ctg gcg tgg gcg atg gac ttt gct gac tgg ccg caa 1248
Asn Arg Ala Arg Leu Ala Trp Ala Met Asp Phe Ala Asp Trp Pro Gln
405 410 415

ctg acc ggg gcg ctg acc gca cat atg acc aat gtg cgc cgg gtg ttt 1296
Leu Thr Gly Ala Leu Thr Ala His Met Thr Asn Val Arg Arg Val Phe
420 425 430

aat gaa ttg att ggc gac gat gaa agt gaa act cag gaa gag tcg ctg 1344
Asn Glu Leu Ile Gly Asp Asp Glu Ser Glu Thr Gln Glu Glu Ser Leu
435 440 445

tcg gaa cag tgg cgt gag ctg tgg cag gat gcg ttg cag gaa gat gac 1392
Ser Glu Gln Trp Arg Glu Leu Trp Gln Asp Ala Leu Gln Glu Asp Asp
450 455 460

act acg cca gtg ctg gcg cat ctt agc gag gat gat cgc aaa cag gtg 1440
Thr Thr Pro Val Leu Ala His Leu Ser Glu Asp Asp Arg Lys Gln Val
465 470 475 480

cta acg ctg att gcc gat ttc cgc aaa gag ctg gat aag cgc acc atc 1488
Leu Thr Leu Ile Ala Asp Phe Arg Lys Glu Leu Asp Lys Arg Thr Ile
485 490 495

ggg ccg cga gga cgt cag gtg ctc gac cat ctg atg ccg cat ctg cta 1536
Gly Pro Arg Gly Arg Gln Val Leu Asp His Leu Met Pro His Leu Leu
500 505 510

agt gat gtc tgt gcg cgt gaa gac gct gcc gtt acg ctg tcg cgc att 1584
Ser Asp Val Cys Ala Arg Glu Asp Ala Ala Val Thr Leu Ser Arg Ile
515 520 525

acc gcc ttg ctg gtg ggg att gtt acc cgc acc acc tat tta gaa ttg 1632
Thr Ala Leu Leu Val Gly Ile Val Thr Arg Thr Thr Tyr Leu Glu Leu
530 535 540

ctc agt gaa ttc ccc gcg gcg ctt aaa cat ttg att tct ctg tgt gcc 1680
Leu Ser Glu Phe Pro Ala Ala Leu Lys His Leu Ile Ser Leu Cys Ala
545 550 555 560

gcg tcg ccg atg att gcc agc cag ctg gcg cgt tat cca tta ttg ctg 1728
Ala Ser Pro Met Ile Ala Ser Gln Leu Ala Arg Tyr Pro Leu Leu Leu
565 570 575

gat gaa ttg ctc gat cca aac acc ctt tac cag ccg acg gcg acc gat 1776
Asp Glu Leu Leu Asp Pro Asn Thr Leu Tyr Gln Pro Thr Ala Thr Asp
580 585 590

gcc tac cgc gat gag ttg cgc cag tat ttg ctg cgc gtg ccg gaa gat 1824
Ala Tyr Arg Asp Glu Leu Arg Gln Tyr Leu Leu Arg Val Pro Glu Asp
595 600 605

gac gaa gag caa cag ctt gag gcg ctg cgt cag ttc aaa cag gcg cag 1872
Asp Glu Glu Gln Gln Leu Glu Ala Leu Arg Gln Phe Lys Gln Ala Gln
610 615 620

ctg tta cgc atc gcc gca gcg gat atc gcc ggt acg cta ccg gtg atg 1920
Leu Leu Arg Ile Ala Ala Ala Asp Ile Ala Gly Thr Leu Pro Val Met
625 630 635 640

aaa gtg agc gat cac tta acc tgg ctg gcg gaa gcc atg ata gat gcc 1968
Lys Val Ser Asp His Leu Thr Trp Leu Ala Glu Ala Met Ile Asp Ala
645 650 655

gtc gtt cag cag gcg tgg gtt caa atg gtt gcc cgc tac ggt aag ccg 2016
Val Val Gln Gln Ala Trp Val Gln Met Val Ala Arg Tyr Gly Lys Pro
660 665 670

aat cac ctg aac gaa cgc gaa ggg cgt ggt ttt gcg gtg gtc ggc tac 2064
Asn His Leu Asn Glu Arg Glu Gly Arg Gly Phe Ala Val Val Gly Tyr
675 680 685

ggc aag ctg ggc ggc tgg gag tta ggc tac agt tcc gat ctt gac ctt 2112
Gly Lys Leu Gly Gly Trp Glu Leu Gly Tyr Ser Ser Asp Leu Asp Leu
690 695 700

atc ttc ctc cat gat tgc cca atg gat gcg atg act gac ggt gag cgg 2160
Ile Phe Leu His Asp Cys Pro Met Asp Ala Met Thr Asp Gly Glu Arg
705 710 715 720

gaa atc gac ggg cgg cag ttt tat ctg cgt ctg gcg caa cgc att atg 2208
Glu Ile Asp Gly Arg Gln Phe Tyr Leu Arg Leu Ala Gln Arg Ile Met
725 730 735

cat ctg ttc agt acg cgt acc tct tcc ggc att ttg tat gaa gtg gat 2256
His Leu Phe Ser Thr Arg Thr Ser Ser Gly Ile Leu Tyr Glu Val Asp
740 745 750

gct cga ctg cgt ccg tcc ggg gcg gcg gga atg ctg gtg aca tcc gca 2304
Ala Arg Leu Arg Pro Ser Gly Ala Ala Gly Met Leu Val Thr Ser Ala
755 760 765

gaa gca ttt gcc gat tat cag aaa aac gag gcc tgg acg tgg gaa cat 2352
Glu Ala Phe Ala Asp Tyr Gln Lys Asn Glu Ala Trp Thr Trp Glu His
770 775 780

cag gcg ctg gtg cgt gcg cgt gta gtg tac ggc gat ccg cag ctc acc 2400
Gln Ala Leu Val Arg Ala Arg Val Val Tyr Gly Asp Pro Gln Leu Thr
785 790 795 800

gcg cac ttt gac gca gtg cgt cgc gag att atg acg ctg ccg cgt gaa 2448
Ala His Phe Asp Ala Val Arg Arg Glu Ile Met Thr Leu Pro Arg Glu
805 810 815

ggt aaa act ctg caa acg gaa gtg cgg gaa atg cgc gag aaa atg cgc 2496
Gly Lys Thr Leu Gln Thr Glu Val Arg Glu Met Arg Glu Lys Met Arg
820 825 830

gct cat ctc ggc aat aaa cat cgc gat cgc ttt gat atc aaa gct gat 2544
Ala His Leu Gly Asn Lys His Arg Asp Arg Phe Asp Ile Lys Ala Asp
835 840 845

gaa ggg gga att acc gat atc gaa ttt att acc caa tat ctg gtg ttg 2592
Glu Gly Gly Ile Thr Asp Ile Glu Phe Ile Thr Gln Tyr Leu Val Leu
850 855 860

cgc tac gct cat gaa aaa ccg aag tta acg cgc tgg tca gac aac gtg 2640

Arg Tyr Ala His Glu Lys Pro Lys Leu Thr Arg Trp Ser Asp Asn Val
865 870 875 880

cgt att ctg gaa cta ctg gcg caa aac gac att atg gaa gag cag gaa 2688
Arg Ile Leu Glu Leu Leu Ala Gln Asn Asp Ile Met Glu Glu Gln Glu
885 890 895

gcg atg gcg ctg acc cgt gct tac act acg ctt cgc gat gaa ctt cat 2736
Ala Met Ala Leu Thr Arg Ala Tyr Thr Thr Leu Arg Asp Glu Leu His
900 905 910

cat ctg gca tta cag gaa ttg ccg ggc cat gtg tcg gag gat tgc ttc 2784
His Leu Ala Leu Gln Glu Leu Pro Gly His Val Ser Glu Asp Cys Phe
915 920 925

acc gca gag cgt gaa ctg gtg cgg gca agc tgg cag aag tgg ctg gtg 2832
Thr Ala Glu Arg Glu Leu Val Arg Ala Ser Trp Gln Lys Trp Leu Val
930 935 940

gaa gaa tga 2841
Glu Glu *
945

<210> 278
<211> 1302
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1302)

<400> 278
atg gct cag gaa atc gaa tta aag ttt att gtt aat cac agt gcc gtt 48
Met Ala Gln Glu Ile Glu Leu Lys Phe Ile Val Asn His Ser Ala Val
1 5 10 15

gag gcg ttg cgt gac cat ctc aat acg ctg ggc ggc gag cac cat gac 96
Glu Ala Leu Arg Asp His Leu Asn Thr Leu Gly Gly Glu His His Asp
20 25 30

ccc gtg cag ttg ctg aat att tac tac gaa acg ccg gat aac tgg ctg 144
Pro Val Gln Leu Leu Asn Ile Tyr Tyr Glu Thr Pro Asp Asn Trp Leu
35 40 45

cgt ggg cac gat atg ggc tta cgt att cgt ggc gaa aac ggt cgc tat 192
Arg Gly His Asp Met Gly Leu Arg Ile Arg Gly Glu Asn Gly Arg Tyr
50 55 60

gag atg acc atg aaa gtt gca gga aga gtg aca ggc ggc tta cat cag 240
Glu Met Thr Met Lys Val Ala Gly Arg Val Thr Gly Gly Leu His Gln
65 70 75 80

cgc ccg gaa tat aac gtg gcg ttg agc gaa ccg acg ctc gac ctg gcg 288
Arg Pro Glu Tyr Asn Val Ala Leu Ser Glu Pro Thr Leu Asp Leu Ala
85 90 95

cag tta ccg acg gaa gtc tgg ccg aac ggc gaa ttg ccc gcc gat ctc 336
Gln Leu Pro Thr Glu Val Trp Pro Asn Gly Glu Leu Pro Ala Asp Leu
100 105 110

```
gcc tcc cgc gtg cag ccg ctg ttc agc acc gat ttt tat cgc gaa aaa     384
Ala Ser Arg Val Gln Pro Leu Phe Ser Thr Asp Phe Tyr Arg Glu Lys
        115                 120                 125

tgg ctg gtg gcg gtc gat ggt agc caa att gaa atc gcc ctc gac cag     432
Trp Leu Val Ala Val Asp Gly Ser Gln Ile Glu Ile Ala Leu Asp Gln
    130                 135                 140

ggg gaa gtg aaa gcg ggt gaa ttt gct gaa cct atc tgt gag ctg gaa     480
Gly Glu Val Lys Ala Gly Glu Phe Ala Glu Pro Ile Cys Glu Leu Glu
145                 150                 155                 160

ctg gaa ctg ctt agc ggc gac acg cgc gcg gtg ctg aaa ctg gcg aac     528
Leu Glu Leu Leu Ser Gly Asp Thr Arg Ala Val Leu Lys Leu Ala Asn
                165                 170                 175

caa ctg gta tcg caa acc gga tta cgc cag ggc agc ctg agc aaa gcg     576
Gln Leu Val Ser Gln Thr Gly Leu Arg Gln Gly Ser Leu Ser Lys Ala
            180                 185                 190

gcg cgt ggc tat cat ctg gcg cag ggc aat ccg gcg cgt gaa atc aaa     624
Ala Arg Gly Tyr His Leu Ala Gln Gly Asn Pro Ala Arg Glu Ile Lys
        195                 200                 205

ccg acc acc att ttg cat gtt gcg gca aaa gcc gat gtg gaa cag ggg     672
Pro Thr Thr Ile Leu His Val Ala Ala Lys Ala Asp Val Glu Gln Gly
    210                 215                 220

ctg gaa gcg gcg ctc gag ctg gcg tta gcg caa tgg cag tat cat gaa     720
Leu Glu Ala Ala Leu Glu Leu Ala Leu Ala Gln Trp Gln Tyr His Glu
225                 230                 235                 240

gaa ctg tgg gta cgc ggc aac gat gcg gcg aaa gaa cag gtg ctg gca     768
Glu Leu Trp Val Arg Gly Asn Asp Ala Ala Lys Glu Gln Val Leu Ala
                245                 250                 255

gcc att agc ctg gtc cgt cat acg ctg atg ctg ttc ggt ggt att gtg     816
Ala Ile Ser Leu Val Arg His Thr Leu Met Leu Phe Gly Gly Ile Val
            260                 265                 270

ccg cgt aaa gcg agc act cac tta cgt gat ctg ctg act caa tgc gag     864
Pro Arg Lys Ala Ser Thr His Leu Arg Asp Leu Leu Thr Gln Cys Glu
        275                 280                 285

gcg acc att gct tct gcg gtg tct gcc gtg acg gcg gtc tac tct acc     912
Ala Thr Ile Ala Ser Ala Val Ser Ala Val Thr Ala Val Tyr Ser Thr
    290                 295                 300

gaa acg gca atg gcg aag ctg gcg ttg acc gaa tgg ttg gta agc aaa     960
Glu Thr Ala Met Ala Lys Leu Ala Leu Thr Glu Trp Leu Val Ser Lys
305                 310                 315                 320

gca tgg cag cca ttt tta gat gcc aaa gcg cag ggc aaa atc agc gac    1008
Ala Trp Gln Pro Phe Leu Asp Ala Lys Ala Gln Gly Lys Ile Ser Asp
                325                 330                 335

tcc ttc aaa cgc ttt gcc gat atc cat ctt tcc cgc cat gcc gct gaa    1056
Ser Phe Lys Arg Phe Ala Asp Ile His Leu Ser Arg His Ala Ala Glu
            340                 345                 350
```

ctg aaa agc gtt ttc tgc cag ccg tta ggc gat cgc tac cgt gac cag 1104
Leu Lys Ser Val Phe Cys Gln Pro Leu Gly Asp Arg Tyr Arg Asp Gln
355 360 365

ttg cca cgc ctg acg cgt gat att gac tca ata ctg ttg ctg gcg ggt 1152
Leu Pro Arg Leu Thr Arg Asp Ile Asp Ser Ile Leu Leu Leu Ala Gly
370 375 380

tac tat gat cct gtc gtc gcg caa gcc tgg ctg gag aac tgg cag ggg 1200
Tyr Tyr Asp Pro Val Val Ala Gln Ala Trp Leu Glu Asn Trp Gln Gly
385 390 395 400

ctg cat cac gct att gcg acc ggg caa cgc atc gaa att gaa cat ttc 1248
Leu His His Ala Ile Ala Thr Gly Gln Arg Ile Glu Ile Glu His Phe
405 410 415

cgt aat gag gca aac aat cag gaa ccg ttc tgg ttg cac agc gga aaa 1296
Arg Asn Glu Ala Asn Asn Gln Glu Pro Phe Trp Leu His Ser Gly Lys
420 425 430

cgt taa 1302
Arg *

<210> 279
<211> 624
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(624)

<400> 279
gtg ctc aac aaa ctc tcc tta ctg ctg aaa gac gca ggt att tcg ctt 48
Met Leu Asn Lys Leu Ser Leu Leu Leu Lys Asp Ala Gly Ile Ser Leu
1 5 10 15

acc gat cac cag aaa aac cag ctt att gcc tac gtg aat atg ctg cat 96
Thr Asp His Gln Lys Asn Gln Leu Ile Ala Tyr Val Asn Met Leu His
20 25 30

aaa tgg aac aaa gcg tac aac ctg act tcg gtc cgc gat cct aat gag 144
Lys Trp Asn Lys Ala Tyr Asn Leu Thr Ser Val Arg Asp Pro Asn Glu
35 40 45

atg ctg gta cgc cat att ctc gat agc att gtg gtg gca ccg tat ctg 192
Met Leu Val Arg His Ile Leu Asp Ser Ile Val Val Ala Pro Tyr Leu
50 55 60

caa ggt gaa cgg ttt atc gat gtc ggc acc gga cca gga ctg cca ggc 240
Gln Gly Glu Arg Phe Ile Asp Val Gly Thr Gly Pro Gly Leu Pro Gly
65 70 75 80

att cca ctc tct atc gtg cgt cct gaa gcc cat ttc act ctg ttg gat 288
Ile Pro Leu Ser Ile Val Arg Pro Glu Ala His Phe Thr Leu Leu Asp
85 90 95

agc ctt ggt aaa cgc gtg cgt ttc ctt cgt cag gtg caa cat gag ctt 336
Ser Leu Gly Lys Arg Val Arg Phe Leu Arg Gln Val Gln His Glu Leu

100 105 110

aaa ctg gag aat att gaa cca gta cag agc agg gta gaa gag ttt cct 384
Lys Leu Glu Asn Ile Glu Pro Val Gln Ser Arg Val Glu Glu Phe Pro
115 120 125

tca gag ccg cca ttt gat ggc gta att agc cgc gct ttt gcc tct ctg 432
Ser Glu Pro Pro Phe Asp Gly Val Ile Ser Arg Ala Phe Ala Ser Leu
130 135 140

aac gat atg gtg agc tgg tgc cac cat ctt cct ggt gag caa ggc cgt 480
Asn Asp Met Val Ser Trp Cys His His Leu Pro Gly Glu Gln Gly Arg
145 150 155 160

ttc tac gcg ctg aaa ggg caa atg ccg gaa gat gaa atc gct ttg ttg 528
Phe Tyr Ala Leu Lys Gly Gln Met Pro Glu Asp Glu Ile Ala Leu Leu
165 170 175

ccc gaa gaa tat cag gtc gaa tca gtg gtt aaa ctt cag gtt cca gcc 576
Pro Glu Glu Tyr Gln Val Glu Ser Val Val Lys Leu Gln Val Pro Ala
180 185 190

ctg gat ggc gaa cgt cat ctg gtg gtg att aaa gca aat aaa att taa 624
Leu Asp Gly Glu Arg His Leu Val Val Ile Lys Ala Asn Lys Ile *
195 200 205

<210> 280
<211> 1890
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(1890)

<400> 280
atg ttt tat ccg gat cct ttt gac gtc atc atc att ggc ggg ggt cat 48
Met Phe Tyr Pro Asp Pro Phe Asp Val Ile Ile Ile Gly Gly Gly His
1 5 10 15

gca ggc acc gag gcc gcg atg gcc gcg gcg cgt atg ggt caa cag act 96
Ala Gly Thr Glu Ala Ala Met Ala Ala Ala Arg Met Gly Gln Gln Thr
20 25 30

ctg ctt ttg aca cac aat atc gac act ctg ggg cag atg agc tgc aac 144
Leu Leu Leu Thr His Asn Ile Asp Thr Leu Gly Gln Met Ser Cys Asn
35 40 45

ccg gcg atc ggc ggt att ggg aag gga cat ctg gta aaa gaa gtg gat 192
Pro Ala Ile Gly Gly Ile Gly Lys Gly His Leu Val Lys Glu Val Asp
50 55 60

gca ctc ggc ggt ctg atg gcg aaa gcg atc gat cag gcg ggt atc cag 240
Ala Leu Gly Gly Leu Met Ala Lys Ala Ile Asp Gln Ala Gly Ile Gln
65 70 75 80

ttt agg ata cta aac gca agc aaa gga ccg gcg gtt cgc gct acc cga 288
Phe Arg Ile Leu Asn Ala Ser Lys Gly Pro Ala Val Arg Ala Thr Arg
85 90 95

```
gct cag gcg gat cgt gtg ctc tac cgt cag gcg gta cgt acg gcg ctg    336
Ala Gln Ala Asp Arg Val Leu Tyr Arg Gln Ala Val Arg Thr Ala Leu
            100                 105                 110

gag aac caa ccg aac ctg atg atc ttc cag cag gcg gtt gaa gat ctt    384
Glu Asn Gln Pro Asn Leu Met Ile Phe Gln Gln Ala Val Glu Asp Leu
        115                 120                 125

att gtc gaa aac gat cgc gtg gtc ggt gct gtt acc caa atg gga ctg    432
Ile Val Glu Asn Asp Arg Val Val Gly Ala Val Thr Gln Met Gly Leu
    130                 135                 140

aag ttc cgt gcc aaa gcc gtc gtg ctc acc gtt ggg acg ttc ctc gac    480
Lys Phe Arg Ala Lys Ala Val Val Leu Thr Val Gly Thr Phe Leu Asp
145                 150                 155                 160

ggt aaa att cat atc ggt ctg gat aat tac agc ggt ggc cgt gct ggt    528
Gly Lys Ile His Ile Gly Leu Asp Asn Tyr Ser Gly Gly Arg Ala Gly
                165                 170                 175

gat ccg ccg tcc att ccg ctt tct cgc cgt ttg cgt gaa ctg ccg ctg    576
Asp Pro Pro Ser Ile Pro Leu Ser Arg Arg Leu Arg Glu Leu Pro Leu
            180                 185                 190

cgc gtt ggt cgt ctg aaa acc ggg aca cca ccg cgt att gat gct cga    624
Arg Val Gly Arg Leu Lys Thr Gly Thr Pro Pro Arg Ile Asp Ala Arg
        195                 200                 205

acc atc gac ttt agc gta ctg gcg caa cag cat ggc gat aac cca atg    672
Thr Ile Asp Phe Ser Val Leu Ala Gln Gln His Gly Asp Asn Pro Met
    210                 215                 220

ccg gta ttc tcg ttt atg ggc aat gcg tcc cag cat ccc cag cag gtg    720
Pro Val Phe Ser Phe Met Gly Asn Ala Ser Gln His Pro Gln Gln Val
225                 230                 235                 240

ccg tgt tat atc act cat acc aac gag aaa acc cat gat gtg atc cgc    768
Pro Cys Tyr Ile Thr His Thr Asn Glu Lys Thr His Asp Val Ile Arg
                245                 250                 255

agt aac ctc gat cgt agc cca atg tac gca ggg gtg atc gaa ggt gtc    816
Ser Asn Leu Asp Arg Ser Pro Met Tyr Ala Gly Val Ile Glu Gly Val
            260                 265                 270

ggc cca cgc tac tgc ccg tcg atc gaa gac aaa gtc atg cgc ttc gcc    864
Gly Pro Arg Tyr Cys Pro Ser Ile Glu Asp Lys Val Met Arg Phe Ala
        275                 280                 285

gac aga aat cag cat cag atc ttc ctt gaa ccg gaa gga ctg acc tct    912
Asp Arg Asn Gln His Gln Ile Phe Leu Glu Pro Glu Gly Leu Thr Ser
    290                 295                 300

aat gaa att tat ccg aac ggt atc tcc acc agc ctg ccg ttc gat gtg    960
Asn Glu Ile Tyr Pro Asn Gly Ile Ser Thr Ser Leu Pro Phe Asp Val
305                 310                 315                 320

cag atg caa atc gtc cgc tct atg cag ggg atg gaa aac gcg aag atc    1008
Gln Met Gln Ile Val Arg Ser Met Gln Gly Met Glu Asn Ala Lys Ile
                325                 330                 335
```

gtg cgt ccg ggt tat gcc att gag tat gac ttc ttc gat cct cgc gac 1056
Val Arg Pro Gly Tyr Ala Ile Glu Tyr Asp Phe Phe Asp Pro Arg Asp
340 345 350

ctg aaa ccg acg ctg gag agc aag ttt atc cag ggg ctg ttc ttt gct 1104
Leu Lys Pro Thr Leu Glu Ser Lys Phe Ile Gln Gly Leu Phe Phe Ala
355 360 365

ggt cag att aac ggc act acc ggt tac gaa gaa gcc gct gcg caa ggt 1152
Gly Gln Ile Asn Gly Thr Thr Gly Tyr Glu Glu Ala Ala Ala Gln Gly
370 375 380

ttg ctg gcc ggt ctt aac gct gcc cgt ctg tct gct gac aaa gaa ggt 1200
Leu Leu Ala Gly Leu Asn Ala Ala Arg Leu Ser Ala Asp Lys Glu Gly
385 390 395 400

tgg gct ccg gca cgt tct cag gcg tat ctc ggc gta cta gtt gat gac 1248
Trp Ala Pro Ala Arg Ser Gln Ala Tyr Leu Gly Val Leu Val Asp Asp
405 410 415

ctg tgc act tta gga acc aaa gaa ccg tat cgt atg ttt act tcg cgc 1296
Leu Cys Thr Leu Gly Thr Lys Glu Pro Tyr Arg Met Phe Thr Ser Arg
420 425 430

gca gaa tat cgt ctg atg cta cgc gaa gat aat gcg gat ctg cgt ttg 1344
Ala Glu Tyr Arg Leu Met Leu Arg Glu Asp Asn Ala Asp Leu Arg Leu
435 440 445

act gaa atc ggt cgt gaa ctg ggc ctg gtg gat gac gaa cgt tgg gcg 1392
Thr Glu Ile Gly Arg Glu Leu Gly Leu Val Asp Asp Glu Arg Trp Ala
450 455 460

cgc ttt aac gag aaa ctt gag aat atc gag cgt gag cgt cag cgt ctg 1440
Arg Phe Asn Glu Lys Leu Glu Asn Ile Glu Arg Glu Arg Gln Arg Leu
465 470 475 480

aaa tcg acc tgg gta acc ccg tcg gcg gaa gct gca gcc gaa gtg aat 1488
Lys Ser Thr Trp Val Thr Pro Ser Ala Glu Ala Ala Ala Glu Val Asn
485 490 495

gct cac ctg act gcg ccg ctt tcc cgt gaa gcc agt ggt gaa gat ctg 1536
Ala His Leu Thr Ala Pro Leu Ser Arg Glu Ala Ser Gly Glu Asp Leu
500 505 510

ctg cgt cgt ccg gaa atg act tat gaa aaa tta acc acg ctg acg ccg 1584
Leu Arg Arg Pro Glu Met Thr Tyr Glu Lys Leu Thr Thr Leu Thr Pro
515 520 525

ttt gcc cct gcg ttg aca gac gaa cag gcg gcg gaa cag gtt gag att 1632
Phe Ala Pro Ala Leu Thr Asp Glu Gln Ala Ala Glu Gln Val Glu Ile
530 535 540

cag gtt aaa tac gaa ggt tat atc gcg cgc cag caa gat gag atc gaa 1680
Gln Val Lys Tyr Glu Gly Tyr Ile Ala Arg Gln Gln Asp Glu Ile Glu
545 550 555 560

aag cag ctg cgt aac gag aac acc ctg cta ccc gcg aca ctg gat tac 1728
Lys Gln Leu Arg Asn Glu Asn Thr Leu Leu Pro Ala Thr Leu Asp Tyr
565 570 575

cgc cag gta tcc ggt ctt tct aac gaa gtg atc gcc aaa ctt aac gat 1776

```
Arg Gln Val Ser Gly Leu Ser Asn Glu Val Ile Ala Lys Leu Asn Asp
            580                 585                 590

cac aaa cca gcc tct atc ggc caa gct tcg cgt att tct ggc gtc acg      1824
His Lys Pro Ala Ser Ile Gly Gln Ala Ser Arg Ile Ser Gly Val Thr
        595                 600                 605

cct gcg gcc atc tcc att ctg ctg gtg tgg ctg aaa aaa cag ggt atg      1872
Pro Ala Ala Ile Ser Ile Leu Leu Val Trp Leu Lys Lys Gln Gly Met
    610                 615                 620

ctg cgt cgt agc gca taa                                              1890
Leu Arg Arg Ser Ala  *
625


<210> 281
<211> 444
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(444)

<400> 281
atg gca gat atc act ctt atc agc ggc agc acc ctc ggc ggt gcc gaa      48
Met Ala Asp Ile Thr Leu Ile Ser Gly Ser Thr Leu Gly Gly Ala Glu
 1               5                  10                  15

tat gta gca gaa cac ctg gct gaa aag ctg gaa gag gcg ggt ttt acc      96
Tyr Val Ala Glu His Leu Ala Glu Lys Leu Glu Glu Ala Gly Phe Thr
            20                  25                  30

acc gaa acg ctg cac ggt ccg ctg tta gaa gat tta cct gcc tca ggg      144
Thr Glu Thr Leu His Gly Pro Leu Leu Glu Asp Leu Pro Ala Ser Gly
        35                  40                  45

atc tgg ctg gtt atc agc tcc acc cac ggt gcc gga gat att ccg gac      192
Ile Trp Leu Val Ile Ser Ser Thr His Gly Ala Gly Asp Ile Pro Asp
    50                  55                  60

aac ctt tct cct ttc tat gaa gca ttg cag gaa cag aag ccc gat ctt      240
Asn Leu Ser Pro Phe Tyr Glu Ala Leu Gln Glu Gln Lys Pro Asp Leu
 65                  70                  75                  80

tct gca gtc cgc ttt ggc gca atc ggt att ggc agt cgt gaa tat gac      288
Ser Ala Val Arg Phe Gly Ala Ile Gly Ile Gly Ser Arg Glu Tyr Asp
                85                  90                  95

acc ttt tgt ggg gct atc gat aaa ctc gag gca gaa ctc aaa aat tcc      336
Thr Phe Cys Gly Ala Ile Asp Lys Leu Glu Ala Glu Leu Lys Asn Ser
            100                 105                 110

ggt gca aaa cag aca ggc gaa aca ctg aag atc aac att ctt gat cac      384
Gly Ala Lys Gln Thr Gly Glu Thr Leu Lys Ile Asn Ile Leu Asp His
        115                 120                 125

gac att ccg gaa gat ccg gca gaa gaa tgg ctg gga tcg tgg gtt aat      432
Asp Ile Pro Glu Asp Pro Ala Glu Glu Trp Leu Gly Ser Trp Val Asn
    130                 135                 140
```

```
tta ctc aaa taa                                                 444
Leu Leu Lys  *
145


<210> 282
<211> 456
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(456)

<400> 282
gtg ggt cag cga aac gtt tcg ctg atg gag aaa aaa atg aaa aaa ggc   48
Met Gly Gln Arg Asn Val Ser Leu Met Glu Lys Lys Met Lys Lys Gly
 1               5                  10                  15

acc gtt ctt aat tct gat att tca tcg gtg atc tcc cgt ctg gga cat   96
Thr Val Leu Asn Ser Asp Ile Ser Ser Val Ile Ser Arg Leu Gly His
            20                  25                  30

acc gat acg ctg gtg gtg tgt gat gct ggt tta ccc atc ccc aaa agt  144
Thr Asp Thr Leu Val Val Cys Asp Ala Gly Leu Pro Ile Pro Lys Ser
        35                  40                  45

aca acg cgt atc gat atg gca tta acc cag ggt gta cct tct ttt atg  192
Thr Thr Arg Ile Asp Met Ala Leu Thr Gln Gly Val Pro Ser Phe Met
    50                  55                  60

cag gtg ctg ggc gtc gtc aca aat gaa atg cag gtc gag gcg gcc att  240
Gln Val Leu Gly Val Val Thr Asn Glu Met Gln Val Glu Ala Ala Ile
 65                  70                  75                  80

atc gcg gaa gag atc aaa cac cat aat ccg caa ctc cac gaa acg ttg  288
Ile Ala Glu Glu Ile Lys His His Asn Pro Gln Leu His Glu Thr Leu
                85                  90                  95

ctc act cac ctt gag cag ctg caa aaa cac cag gga aat acc att gaa  336
Leu Thr His Leu Glu Gln Leu Gln Lys His Gln Gly Asn Thr Ile Glu
            100                 105                 110

att cgt tac acc acg cat gaa caa ttc aaa caa caa acc gca gaa agt  384
Ile Arg Tyr Thr Thr His Glu Gln Phe Lys Gln Gln Thr Ala Glu Ser
        115                 120                 125

cag gcg gta att cgc agc gga gaa tgt tct ccg tat gcg aat atc att  432
Gln Ala Val Ile Arg Ser Gly Glu Cys Ser Pro Tyr Ala Asn Ile Ile
    130                 135                 140

ctc tgt gct ggc gtg acg ttc tga                                  456
Leu Cys Ala Gly Val Thr Phe  *
145                 150


<210> 283
<211> 1506
<212> DNA
<213> Escherichia coli
```

<220>
<221> CDS
<222> (1)...(1506)

<400> 283
```
atg gaa gca tta ctt cag ctt aaa ggc atc gat aaa gcc ttc ccg ggc    48
Met Glu Ala Leu Leu Gln Leu Lys Gly Ile Asp Lys Ala Phe Pro Gly
 1               5                  10                  15

gta aaa gcc ctc tcg ggc gca gcg tta aat gtc tat ccg ggc cgc gtg    96
Val Lys Ala Leu Ser Gly Ala Ala Leu Asn Val Tyr Pro Gly Arg Val
            20                  25                  30

atg gcg ctg gtg ggc gaa aac ggc gcg ggt aaa tcc acc atg atg aaa   144
Met Ala Leu Val Gly Glu Asn Gly Ala Gly Lys Ser Thr Met Met Lys
        35                  40                  45

gtg ctt act ggc atc tat act cgc gat gcc ggt acg ctt tta tgg ctg   192
Val Leu Thr Gly Ile Tyr Thr Arg Asp Ala Gly Thr Leu Leu Trp Leu
    50                  55                  60

ggg aaa gaa acg aca ttt acc ggg cca aaa tct tcc cag gaa gcc ggg   240
Gly Lys Glu Thr Thr Phe Thr Gly Pro Lys Ser Ser Gln Glu Ala Gly
65                  70                  75                  80

att ggg att atc cat cag gaa ctg aac ctg atc ccg cag ttg acc att   288
Ile Gly Ile Ile His Gln Glu Leu Asn Leu Ile Pro Gln Leu Thr Ile
                85                  90                  95

gcc gaa aac att ttc ctc ggt cgt gag ttt gtt aat cgc ttt ggc aaa   336
Ala Glu Asn Ile Phe Leu Gly Arg Glu Phe Val Asn Arg Phe Gly Lys
            100                 105                 110

att gac tgg aaa acc atg tat gcc gaa gcg gat aaa ttg ctg gct aaa   384
Ile Asp Trp Lys Thr Met Tyr Ala Glu Ala Asp Lys Leu Leu Ala Lys
        115                 120                 125

ctt aac ctg cgc ttt aaa agc gac aag ctg gtg ggc gat ctt tcc atc   432
Leu Asn Leu Arg Phe Lys Ser Asp Lys Leu Val Gly Asp Leu Ser Ile
    130                 135                 140

ggt gac cag caa atg gtt gaa atc gcc aaa gtg ctg agc ttt gag tcg   480
Gly Asp Gln Gln Met Val Glu Ile Ala Lys Val Leu Ser Phe Glu Ser
145                 150                 155                 160

aaa gtc atc att atg gat gaa ccg acc gat gcg ctg acc gat acc gaa   528
Lys Val Ile Ile Met Asp Glu Pro Thr Asp Ala Leu Thr Asp Thr Glu
                165                 170                 175

acc gaa tcc ctg ttc cgc gtc atc cgc gag ctg aaa tcg caa ggc cgc   576
Thr Glu Ser Leu Phe Arg Val Ile Arg Glu Leu Lys Ser Gln Gly Arg
            180                 185                 190

ggt att gtc tat atc tcc cac cgc atg aaa gaa atc ttc gag att tgc   624
Gly Ile Val Tyr Ile Ser His Arg Met Lys Glu Ile Phe Glu Ile Cys
        195                 200                 205

gat gac gtt acc gtt ttt cgt gat ggg caa ttt att gct gag cgc gaa   672
Asp Asp Val Thr Val Phe Arg Asp Gly Gln Phe Ile Ala Glu Arg Glu
    210                 215                 220
```

gtg gca tca ctg acc gaa gat tcg ctg att gag atg atg gtg ggt cgc 720
Val Ala Ser Leu Thr Glu Asp Ser Leu Ile Glu Met Met Val Gly Arg
225 230 235 240

aag ctg gaa gat caa tat ccg cac ctg gac aaa gcg ccg gga gat atc 768
Lys Leu Glu Asp Gln Tyr Pro His Leu Asp Lys Ala Pro Gly Asp Ile
245 250 255

cgc ctg aaa gtc gat aat ctc tgc gga cct ggc gtt aac gat gtc tct 816
Arg Leu Lys Val Asp Asn Leu Cys Gly Pro Gly Val Asn Asp Val Ser
260 265 270

ttt act tta cgc aaa ggc gaa att ctt ggc gtc tct ggt ttg atg ggc 864
Phe Thr Leu Arg Lys Gly Glu Ile Leu Gly Val Ser Gly Leu Met Gly
275 280 285

gca ggt cgt acc gaa ctg atg aaa gtg ctc tac ggc gca cta ccg cgc 912
Ala Gly Arg Thr Glu Leu Met Lys Val Leu Tyr Gly Ala Leu Pro Arg
290 295 300

acc agc ggt tac gtc acc ctg gat ggg cat gaa gtc gtt acc cgt tca 960
Thr Ser Gly Tyr Val Thr Leu Asp Gly His Glu Val Val Thr Arg Ser
305 310 315 320

ccg cag gat ggc ctg gca aac ggc att gtg tat atc tcc gaa gac cgt 1008
Pro Gln Asp Gly Leu Ala Asn Gly Ile Val Tyr Ile Ser Glu Asp Arg
325 330 335

aaa cgt gac ggt tta gtg ttg ggc atg tca gta aaa gag aac atg tcg 1056
Lys Arg Asp Gly Leu Val Leu Gly Met Ser Val Lys Glu Asn Met Ser
340 345 350

ctg aca gcg ctg cgc tac ttc agc cgc gct ggc ggc agt ttg aag cat 1104
Leu Thr Ala Leu Arg Tyr Phe Ser Arg Ala Gly Gly Ser Leu Lys His
355 360 365

gcc gat gaa cag cag gct gtg agt gat ttc att cgt ctg ttt aat gtg 1152
Ala Asp Glu Gln Gln Ala Val Ser Asp Phe Ile Arg Leu Phe Asn Val
370 375 380

aaa act ccg tcg atg gaa cag gca att ggt ctg ctt tcc ggt ggc aat 1200
Lys Thr Pro Ser Met Glu Gln Ala Ile Gly Leu Leu Ser Gly Gly Asn
385 390 395 400

cag caa aaa gtg gcg att gcc cgc ggt ctg atg aca cgc ccc aaa gtg 1248
Gln Gln Lys Val Ala Ile Ala Arg Gly Leu Met Thr Arg Pro Lys Val
405 410 415

ttg atc ctt gat gaa cct acc cgt ggc gta gat gtc ggc gcg aaa aaa 1296
Leu Ile Leu Asp Glu Pro Thr Arg Gly Val Asp Val Gly Ala Lys Lys
420 425 430

gag atc tat caa ctg att aac cag ttc aaa gcc gat ggc ttg agc atc 1344
Glu Ile Tyr Gln Leu Ile Asn Gln Phe Lys Ala Asp Gly Leu Ser Ile
435 440 445

att ctg gtg tca tcg gag atg cca gaa gta tta ggc atg agc gat cgc 1392
Ile Leu Val Ser Ser Glu Met Pro Glu Val Leu Gly Met Ser Asp Arg
450 455 460

atc atc gtc atg cat gaa ggg cat ctc agc ggg gaa ttt act cgt gag 1440
Ile Ile Val Met His Glu Gly His Leu Ser Gly Glu Phe Thr Arg Glu
465 470 475 480

cag gcc acc cag gaa gtg tta atg gct gcc gct gtg ggc aag ctt aat 1488
Gln Ala Thr Gln Glu Val Leu Met Ala Ala Ala Val Gly Lys Leu Asn
485 490 495

cgc gtg aat cag gag taa 1506
Arg Val Asn Gln Glu *
500

<210> 284
<211> 966
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(966)

<400> 284
atg aca acc cag act gtc tct ggt cgc cgt tat ttc acg aaa gcg tgg 48
Met Thr Thr Gln Thr Val Ser Gly Arg Arg Tyr Phe Thr Lys Ala Trp
1 5 10 15

ctg atg gag cag aaa tcg ctt atc gct ctg ctg gtg ctg atc gcg att 96
Leu Met Glu Gln Lys Ser Leu Ile Ala Leu Leu Val Leu Ile Ala Ile
20 25 30

gtc tcg acg tta agc ccg aac ttt ttc acc atc aat aac tta ttc aat 144
Val Ser Thr Leu Ser Pro Asn Phe Phe Thr Ile Asn Asn Leu Phe Asn
35 40 45

att ctc cag caa acc tca gtg aac gcc att atg gcg gtc ggg atg acg 192
Ile Leu Gln Gln Thr Ser Val Asn Ala Ile Met Ala Val Gly Met Thr
50 55 60

ctg gtg atc ctg acg tcg ggc atc gac tta tcg gta ggt tct ctg ttg 240
Leu Val Ile Leu Thr Ser Gly Ile Asp Leu Ser Val Gly Ser Leu Leu
65 70 75 80

gcg ctg acc ggc gca gtt gct gca tct atc gtc ggc att gaa gtc aat 288
Ala Leu Thr Gly Ala Val Ala Ala Ser Ile Val Gly Ile Glu Val Asn
85 90 95

gcg ctg gtg gct gtc gct gct gct ctc gcg tta ggt gcc gca att ggt 336
Ala Leu Val Ala Val Ala Ala Ala Leu Ala Leu Gly Ala Ala Ile Gly
100 105 110

gcg gta acc ggg gtg att gta gcg aaa ggt cgc gtc cag gcg ttt atc 384
Ala Val Thr Gly Val Ile Val Ala Lys Gly Arg Val Gln Ala Phe Ile
115 120 125

gct acg ctg gtt atg atg ctt tta ctg cgc ggc gtg acc atg gtt tat 432
Ala Thr Leu Val Met Met Leu Leu Leu Arg Gly Val Thr Met Val Tyr
130 135 140

acc aac ggt agc cca gtg aat acc ggc ttt act gag aac gcc gat ctg 480
Thr Asn Gly Ser Pro Val Asn Thr Gly Phe Thr Glu Asn Ala Asp Leu

145 150 155 160

ttt ggc tgg ttt ggt att ggt cgt ccg ctg ggc gta ccg acg cca gtc 528
Phe Gly Trp Phe Gly Ile Gly Arg Pro Leu Gly Val Pro Thr Pro Val
165 170 175

tgg atc atg ggg att gtc ttc ctc gcg gcc tgg tac atg ctg cat cac 576
Trp Ile Met Gly Ile Val Phe Leu Ala Ala Trp Tyr Met Leu His His
180 185 190

acg cgt ctg ggg cgt tac atc tac gcg ctg ggc ggc aac gaa gcg gca 624
Thr Arg Leu Gly Arg Tyr Ile Tyr Ala Leu Gly Gly Asn Glu Ala Ala
195 200 205

acg cgt ctt tct ggt atc aac gtc aat aaa atc aaa atc atc gtc tat 672
Thr Arg Leu Ser Gly Ile Asn Val Asn Lys Ile Lys Ile Ile Val Tyr
210 215 220

tct ctt tgt ggt ctg ctg gca tcg ctg gcc ggg atc att gaa gtg gcg 720
Ser Leu Cys Gly Leu Leu Ala Ser Leu Ala Gly Ile Ile Glu Val Ala
225 230 235 240

cgt ctc tcc tcc gca caa ccc acg gcg ggg act ggc tat gag ctg gat 768
Arg Leu Ser Ser Ala Gln Pro Thr Ala Gly Thr Gly Tyr Glu Leu Asp
245 250 255

gct att gct gcg gtg gtt ctg ggc ggt acg agt ctg gcg ggc gga aaa 816
Ala Ile Ala Ala Val Val Leu Gly Gly Thr Ser Leu Ala Gly Gly Lys
260 265 270

ggt cgc att gtt ggg acg ttg atc ggc gca tta att ctt ggc ttc ctt 864
Gly Arg Ile Val Gly Thr Leu Ile Gly Ala Leu Ile Leu Gly Phe Leu
275 280 285

aat aat gga ttg aat ttg tta ggt gtt tcc tcc tat tac cag atg atc 912
Asn Asn Gly Leu Asn Leu Leu Gly Val Ser Ser Tyr Tyr Gln Met Ile
290 295 300

gtc aaa gcg gtg gtg att ttg ctg gcg gtg ctg gta gac aac aaa aag 960
Val Lys Ala Val Val Ile Leu Leu Ala Val Leu Val Asp Asn Lys Lys
305 310 315 320

cag taa 966
Gln *

<210> 285
<211> 891
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(891)

<400> 285
atg aac atg aaa aaa ctg gct acc ctg gtt tcc gct gtt gcg cta agc 48
Met Asn Met Lys Lys Leu Ala Thr Leu Val Ser Ala Val Ala Leu Ser
1 5 10 15

gcc acc gtc agt gcg aat gcg atg gca aaa gac acc atc gcg ctg gtg 96
Ala Thr Val Ser Ala Asn Ala Met Ala Lys Asp Thr Ile Ala Leu Val
20 25 30

gtc tcc acg ctt aac aac ccg ttc ttt gta tcg ctg aaa gat ggc gcg 144
Val Ser Thr Leu Asn Asn Pro Phe Phe Val Ser Leu Lys Asp Gly Ala
35 40 45

cag aaa gag gcg gat aaa ctt ggc tat aac ctg gtg gtg ctg gac tcc 192
Gln Lys Glu Ala Asp Lys Leu Gly Tyr Asn Leu Val Val Leu Asp Ser
50 55 60

cag aac aac ccg gcg aaa gag ctg gcg aac gtg cag gac tta acc gtt 240
Gln Asn Asn Pro Ala Lys Glu Leu Ala Asn Val Gln Asp Leu Thr Val
65 70 75 80

cgc ggc aca aaa att ctg ctg att aac ccg acc gac tcc gac gca gtg 288
Arg Gly Thr Lys Ile Leu Leu Ile Asn Pro Thr Asp Ser Asp Ala Val
85 90 95

ggt aat gct gtg aag atg gct aac cag gcg aac atc ccg gtt atc act 336
Gly Asn Ala Val Lys Met Ala Asn Gln Ala Asn Ile Pro Val Ile Thr
100 105 110

ctt gac cgc cag gca acg aaa ggt gaa gtg gtg agc cac att gct tct 384
Leu Asp Arg Gln Ala Thr Lys Gly Glu Val Val Ser His Ile Ala Ser
115 120 125

gat aac gta ctg ggc ggc aaa atc gct ggt gat tac atc gcg aag aaa 432
Asp Asn Val Leu Gly Gly Lys Ile Ala Gly Asp Tyr Ile Ala Lys Lys
130 135 140

gcg ggt gaa ggt gcc aaa gtt atc gag ctg caa ggc att gct ggt aca 480
Ala Gly Glu Gly Ala Lys Val Ile Glu Leu Gln Gly Ile Ala Gly Thr
145 150 155 160

tcc gca gcc cgt gaa cgt ggc gaa ggc ttc cag cag gcc gtt gct gct 528
Ser Ala Ala Arg Glu Arg Gly Glu Gly Phe Gln Gln Ala Val Ala Ala
165 170 175

cac aag ttt aat gtt ctt gcc agc cag cca gca gat ttt gat cgc att 576
His Lys Phe Asn Val Leu Ala Ser Gln Pro Ala Asp Phe Asp Arg Ile
180 185 190

aaa ggt ttg aac gta atg cag aac ctg ttg acc gct cat ccg gat gtt 624
Lys Gly Leu Asn Val Met Gln Asn Leu Leu Thr Ala His Pro Asp Val
195 200 205

cag gct gta ttc gcg cag aat gat gaa atg gcg ctg ggc gcg ctg cgc 672
Gln Ala Val Phe Ala Gln Asn Asp Glu Met Ala Leu Gly Ala Leu Arg
210 215 220

gca ctg caa act gcc ggt aaa tcg gat gtg atg gtc gtc gga ttt gac 720
Ala Leu Gln Thr Ala Gly Lys Ser Asp Val Met Val Val Gly Phe Asp
225 230 235 240

ggt aca ccg gat ggc gaa aaa gcg gtg aat gat ggc aaa cta gca gcg 768
Gly Thr Pro Asp Gly Glu Lys Ala Val Asn Asp Gly Lys Leu Ala Ala
245 250 255

act atc gct cag cta ccc gat cag att ggc gcg aaa ggc gtc gaa acc 816

```
Thr Ile Ala Gln Leu Pro Asp Gln Ile Gly Ala Lys Gly Val Glu Thr
            260                 265                 270

gca gat aaa gtg ctg aaa ggc gag aaa gtt cag gct aag tat ccg gtt      864
Ala Asp Lys Val Leu Lys Gly Glu Lys Val Gln Ala Lys Tyr Pro Val
        275                 280                 285

gat ctg aaa ctg gtt gtt aag cag tag                                  891
Asp Leu Lys Leu Val Val Lys Gln  *
    290                 295


<210> 286
<211> 930
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)...(930)

<400> 286
atg caa aac gca ggc agc ctc gtt gtt ctt ggc agc att aat gct gac      48
Met Gln Asn Ala Gly Ser Leu Val Val Leu Gly Ser Ile Asn Ala Asp
 1               5                  10                  15

cac att ctt aat ctt caa tct ttt cct act cca ggc gaa acc gta acc      96
His Ile Leu Asn Leu Gln Ser Phe Pro Thr Pro Gly Glu Thr Val Thr
            20                  25                  30

ggt aac cac tat cag gtt gca ttt ggc ggc aaa ggc gcg aat cag gct      144
Gly Asn His Tyr Gln Val Ala Phe Gly Gly Lys Gly Ala Asn Gln Ala
        35                  40                  45

gtg gct gct ggg cgt agc ggt gcg aat atc gcg ttt att gcc tgt acg      192
Val Ala Ala Gly Arg Ser Gly Ala Asn Ile Ala Phe Ile Ala Cys Thr
    50                  55                  60

ggt gat gac agc att ggt gag agc gtt cgc cag cag ctc gcc act gat      240
Gly Asp Asp Ser Ile Gly Glu Ser Val Arg Gln Gln Leu Ala Thr Asp
65                  70                  75                  80

aac att gat att act ccg gtc agc gtg atc aaa ggc gaa tca aca ggt      288
Asn Ile Asp Ile Thr Pro Val Ser Val Ile Lys Gly Glu Ser Thr Gly
                85                  90                  95

gtg gcg ctg att ttt gtt aat ggc gaa ggt gag aat gtc atc ggt att      336
Val Ala Leu Ile Phe Val Asn Gly Glu Gly Glu Asn Val Ile Gly Ile
            100                 105                 110

cat gcc ggc gct aat gct gcc ctt tcc ccg gcg ctg gtg gaa gcg caa      384
His Ala Gly Ala Asn Ala Ala Leu Ser Pro Ala Leu Val Glu Ala Gln
        115                 120                 125

cgt gag cgt att gcc aac gcg tca gca tta tta atg cag ctg gaa tca      432
Arg Glu Arg Ile Ala Asn Ala Ser Ala Leu Leu Met Gln Leu Glu Ser
    130                 135                 140

cca ctc gaa agt gtg atg gca gcg gcg aaa atc gcc cat caa aat aag      480
Pro Leu Glu Ser Val Met Ala Ala Ala Lys Ile Ala His Gln Asn Lys
145                 150                 155                 160
```

act atc gtt gcg ctt aac ccg gct ccg gct cgc gaa ctt cct gac gaa 528
Thr Ile Val Ala Leu Asn Pro Ala Pro Ala Arg Glu Leu Pro Asp Glu
165 170 175

ctg ctg gcg ctg gtg gac att att acg cca aac gaa acg gaa gca gaa 576
Leu Leu Ala Leu Val Asp Ile Ile Thr Pro Asn Glu Thr Glu Ala Glu
180 185 190

aag ctc acc ggt att cgt gtt gaa aat gat gaa gat gca gcg aag gcg 624
Lys Leu Thr Gly Ile Arg Val Glu Asn Asp Glu Asp Ala Ala Lys Ala
195 200 205

gcg cag gta ctg cat gaa aaa ggt atc cgt act gta ctg att act tta 672
Ala Gln Val Leu His Glu Lys Gly Ile Arg Thr Val Leu Ile Thr Leu
210 215 220

gga agt cgt ggt gta tgg gct agc gtg aat ggt gaa ggt cag cgc gtt 720
Gly Ser Arg Gly Val Trp Ala Ser Val Asn Gly Glu Gly Gln Arg Val
225 230 235 240

cct gga ttc cgg gtg cag gct gtc gat acc att gct gcc gga gat acc 768
Pro Gly Phe Arg Val Gln Ala Val Asp Thr Ile Ala Ala Gly Asp Thr
245 250 255

ttt aac ggt gcg tta atc acg gca ttg ctg gaa gaa aaa cca ttg cca 816
Phe Asn Gly Ala Leu Ile Thr Ala Leu Leu Glu Glu Lys Pro Leu Pro
260 265 270

gag gcg att cgt ttt gcc cat gct gcc gct gcg att gcc gta aca cgt 864
Glu Ala Ile Arg Phe Ala His Ala Ala Ala Ala Ile Ala Val Thr Arg
275 280 285

aaa ggc gca caa cct tcc gta ccg tgg cgt gaa gag atc gac gca ttt 912
Lys Gly Ala Gln Pro Ser Val Pro Trp Arg Glu Glu Ile Asp Ala Phe
290 295 300

tta gac agg cag agg tga 930
Leu Asp Arg Gln Arg *
305

<210> 287
<211> 38
<212> PRT
<213> Escherichia coli

<400> 287
Met Lys Val Arg Ala Ser Val Lys Lys Leu Cys Arg Asn Cys Lys Ile
1 5 10 15
Val Lys Arg Asp Gly Val Ile Arg Val Ile Cys Ser Ala Glu Pro Lys
20 25 30
His Lys Gln Arg Gln Gly
35

<210> 288
<211> 443
<212> PRT
<213> Escherichia coli

<400> 288
Met Ala Lys Gln Pro Gly Leu Asp Phe Gln Ser Ala Lys Gly Gly Leu
1 5 10 15
Gly Glu Leu Lys Arg Arg Leu Leu Phe Val Ile Gly Ala Leu Ile Val
20 25 30
Phe Arg Ile Gly Ser Phe Ile Pro Ile Pro Gly Ile Asp Ala Ala Val
35 40 45
Leu Ala Lys Leu Leu Glu Gln Gln Arg Gly Thr Ile Ile Glu Met Phe
50 55 60
Asn Met Phe Ser Gly Gly Ala Leu Ser Arg Ala Ser Ile Phe Ala Leu
65 70 75 80
Gly Ile Met Pro Tyr Ile Ser Ala Ser Ile Ile Ile Gln Leu Leu Thr
85 90 95
Val Val His Pro Thr Leu Ala Glu Ile Lys Lys Glu Gly Glu Ser Gly
100 105 110
Arg Arg Lys Ile Ser Gln Tyr Thr Arg Tyr Gly Thr Leu Val Leu Ala
115 120 125
Ile Phe Gln Ser Ile Gly Ile Ala Thr Gly Leu Pro Asn Met Pro Gly
130 135 140
Met Gln Gly Leu Val Ile Asn Pro Gly Phe Ala Phe Tyr Phe Thr Ala
145 150 155 160
Val Val Ser Leu Val Thr Gly Thr Met Phe Leu Met Trp Leu Gly Glu
165 170 175
Gln Ile Thr Glu Arg Gly Ile Gly Asn Gly Ile Ser Ile Ile Ile Phe
180 185 190
Ala Gly Ile Val Ala Gly Leu Pro Pro Ala Ile Ala His Thr Ile Glu
195 200 205
Gln Ala Arg Gln Gly Asp Leu His Phe Leu Val Leu Leu Leu Val Ala
210 215 220
Val Leu Val Phe Ala Val Thr Phe Phe Val Val Phe Val Glu Arg Gly
225 230 235 240
Gln Arg Arg Ile Val Val Asn Tyr Ala Lys Arg Gln Gln Gly Arg Arg
245 250 255
Val Tyr Ala Ala Gln Ser Thr His Leu Pro Leu Lys Val Asn Met Ala
260 265 270
Gly Val Ile Pro Ala Ile Phe Ala Ser Ser Ile Ile Leu Phe Pro Ala
275 280 285
Thr Ile Ala Ser Trp Phe Gly Gly Gly Thr Gly Trp Asn Trp Leu Thr
290 295 300
Thr Ile Ser Leu Tyr Leu Gln Pro Gly Gln Pro Leu Tyr Val Leu Leu
305 310 315 320
Tyr Ala Ser Ala Ile Ile Phe Phe Cys Phe Phe Tyr Thr Ala Leu Val
325 330 335
Phe Asn Pro Arg Glu Thr Ala Asp Asn Leu Lys Lys Ser Gly Ala Phe
340 345 350
Val Pro Gly Ile Arg Pro Gly Glu Gln Thr Ala Lys Tyr Ile Asp Lys
355 360 365
Val Met Thr Arg Leu Thr Leu Val Gly Ala Leu Tyr Ile Thr Phe Ile
370 375 380
Cys Leu Ile Pro Glu Phe Met Arg Asp Ala Met Lys Val Pro Phe Tyr
385 390 395 400
Phe Gly Gly Thr Ser Leu Leu Ile Val Val Val Val Ile Met Asp Phe
405 410 415
Met Ala Gln Val Gln Thr Leu Met Met Ser Ser Gln Tyr Glu Ser Ala
420 425 430
Leu Lys Lys Ala Asn Leu Lys Gly Tyr Gly Arg
435 440

<210> 289
<211> 144

<212> PRT
<213> Escherichia coli

<400> 289
Met Arg Leu Asn Thr Leu Ser Pro Ala Glu Gly Ser Lys Lys Ala Gly
1 5 10 15
Lys Arg Leu Gly Arg Gly Ile Gly Ser Gly Leu Gly Lys Thr Gly Gly
20 25 30
Arg Gly His Lys Gly Gln Lys Ser Arg Ser Gly Gly Gly Val Arg Arg
35 40 45
Gly Phe Glu Gly Gly Gln Met Pro Leu Tyr Arg Arg Leu Pro Lys Phe
50 55 60
Gly Phe Thr Ser Arg Lys Ala Ala Ile Thr Ala Glu Ile Arg Leu Ser
65 70 75 80
Asp Leu Ala Lys Val Glu Gly Gly Val Val Asp Leu Asn Thr Leu Lys
85 90 95
Ala Ala Asn Ile Ile Gly Ile Gln Ile Glu Phe Ala Lys Val Ile Leu
100 105 110
Ala Gly Glu Val Thr Thr Pro Val Thr Val Arg Gly Leu Arg Val Thr
115 120 125
Lys Gly Ala Arg Ala Ala Ile Glu Ala Ala Gly Gly Lys Ile Glu Glu
130 135 140

<210> 290
<211> 59
<212> PRT
<213> Escherichia coli

<400> 290
Met Ala Lys Thr Ile Lys Ile Thr Gln Thr Arg Ser Ala Ile Gly Arg
1 5 10 15
Leu Pro Lys His Lys Ala Thr Leu Leu Gly Leu Gly Leu Arg Arg Ile
20 25 30
Gly His Thr Val Glu Arg Glu Asp Thr Pro Ala Ile Arg Gly Met Ile
35 40 45
Asn Ala Val Ser Phe Met Val Lys Val Glu Glu
50 55

<210> 291
<211> 167
<212> PRT
<213> Escherichia coli

<400> 291
Met Ala His Ile Glu Lys Gln Ala Gly Glu Leu Gln Glu Lys Leu Ile
1 5 10 15
Ala Val Asn Arg Val Ser Lys Thr Val Lys Gly Gly Arg Ile Phe Ser
20 25 30
Phe Thr Ala Leu Thr Val Val Gly Asp Gly Asn Gly Arg Val Gly Phe
35 40 45
Gly Tyr Gly Lys Ala Arg Glu Val Pro Ala Ala Ile Gln Lys Ala Met
50 55 60
Glu Lys Ala Arg Arg Asn Met Ile Asn Val Ala Leu Asn Asn Gly Thr
65 70 75 80
Leu Gln His Pro Val Lys Gly Val His Thr Gly Ser Arg Val Phe Met
85 90 95
Gln Pro Ala Ser Glu Gly Thr Gly Ile Ile Ala Gly Gly Ala Met Arg
100 105 110
Ala Val Leu Glu Val Ala Gly Val His Asn Val Leu Ala Lys Ala Tyr

```
        115                 120                 125
Gly Ser Thr Asn Pro Ile Asn Val Val Arg Ala Thr Ile Asp Gly Leu
    130                 135                 140
Glu Asn Met Asn Ser Pro Glu Met Val Ala Ala Lys Arg Gly Lys Ser
145                 150                 155                 160
Val Glu Glu Ile Leu Gly Lys
                165
```

<210> 292
<211> 117
<212> PRT
<213> Escherichia coli

<400> 292

```
Met Asp Lys Lys Ser Ala Arg Ile Arg Arg Ala Thr Arg Ala Arg Arg
 1               5                  10                  15
Lys Leu Gln Glu Leu Gly Ala Thr Arg Leu Val Val His Arg Thr Pro
            20                  25                  30
Arg His Ile Tyr Ala Gln Val Ile Ala Pro Asn Gly Ser Glu Val Leu
        35                  40                  45
Val Ala Ala Ser Thr Val Glu Lys Ala Ile Ala Glu Gln Leu Lys Tyr
    50                  55                  60
Thr Gly Asn Lys Asp Ala Ala Ala Ala Val Gly Lys Ala Val Ala Glu
65                  70                  75                  80
Arg Ala Leu Glu Lys Gly Ile Lys Asp Val Ser Phe Asp Arg Ser Gly
                85                  90                  95
Phe Gln Tyr His Gly Arg Val Gln Ala Leu Ala Asp Ala Ala Arg Glu
            100                 105                 110
Ala Gly Leu Gln Phe
        115
```

<210> 293
<211> 177
<212> PRT
<213> Escherichia coli

<400> 293

```
Met Ser Arg Val Ala Lys Ala Pro Val Val Val Pro Ala Gly Val Asp
 1               5                  10                  15
Val Lys Ile Asn Gly Gln Val Ile Thr Ile Lys Gly Lys Asn Gly Glu
            20                  25                  30
Leu Thr Arg Thr Leu Asn Asp Ala Val Glu Val Lys His Ala Asp Asn
        35                  40                  45
Thr Leu Thr Phe Gly Pro Arg Asp Gly Tyr Ala Asp Gly Trp Ala Gln
    50                  55                  60
Ala Gly Thr Ala Arg Ala Leu Leu Asn Ser Met Val Ile Gly Val Thr
65                  70                  75                  80
Glu Gly Phe Thr Lys Lys Leu Gln Leu Val Gly Val Gly Tyr Arg Ala
                85                  90                  95
Ala Val Lys Gly Asn Val Ile Asn Leu Ser Leu Gly Phe Ser His Pro
            100                 105                 110
Val Asp His Gln Leu Pro Ala Gly Ile Thr Ala Glu Cys Pro Thr Gln
        115                 120                 125
Thr Glu Ile Val Leu Lys Gly Ala Asp Lys Gln Val Ile Gly Gln Val
    130                 135                 140
Ala Ala Asp Leu Arg Ala Tyr Arg Arg Pro Glu Pro Tyr Lys Gly Lys
145                 150                 155                 160
Gly Val Arg Tyr Ala Asp Glu Val Val Arg Thr Lys Glu Ala Lys Lys
                165                 170                 175
```

Lys

<210> 294
<211> 130
<212> PRT
<213> Escherichia coli

<400> 294
Met Ser Met Gln Asp Pro Ile Ala Asp Met Leu Thr Arg Ile Arg Asn
1   5   10   15
Gly Gln Ala Ala Asn Lys Ala Ala Val Thr Met Pro Ser Ser Lys Leu
            20   25   30
Lys Val Ala Ile Ala Asn Val Leu Lys Glu Glu Gly Phe Ile Glu Asp
        35   40   45
Phe Lys Val Glu Gly Asp Thr Lys Pro Glu Leu Glu Leu Thr Leu Lys
    50   55   60
Tyr Phe Gln Gly Lys Ala Val Val Glu Ser Ile Gln Arg Val Ser Arg
65   70   75   80
Pro Gly Leu Arg Ile Tyr Lys Arg Lys Asp Glu Leu Pro Lys Val Met
                85   90   95
Ala Gly Leu Gly Ile Ala Val Val Ser Thr Ser Lys Gly Val Met Thr
            100   105   110
Asp Arg Ala Ala Arg Gln Ala Gly Leu Gly Gly Glu Ile Ile Cys Tyr
        115   120   125
Val Ala
    130

<210> 295
<211> 101
<212> PRT
<213> Escherichia coli

<400> 295
Met Ala Lys Gln Ser Met Lys Ala Arg Glu Val Lys Arg Val Ala Leu
1   5   10   15
Ala Asp Lys Tyr Phe Ala Lys Arg Ala Glu Leu Lys Ala Ile Ile Ser
            20   25   30
Asp Val Asn Ala Ser Asp Glu Asp Arg Trp Asn Ala Val Leu Lys Leu
        35   40   45
Gln Thr Leu Pro Arg Asp Ser Ser Pro Ser Arg Gln Arg Asn Arg Cys
    50   55   60
Arg Gln Thr Gly Arg Pro His Gly Phe Leu Arg Lys Phe Gly Leu Ser
65   70   75   80
Arg Ile Lys Val Arg Glu Ala Ala Met Arg Gly Glu Ile Pro Gly Leu
                85   90   95
Lys Lys Ala Ser Trp
            100

<210> 296
<211> 179
<212> PRT
<213> Escherichia coli

<400> 296
Met Ala Lys Leu His Asp Tyr Tyr Lys Asp Glu Val Val Lys Lys Leu
1   5   10   15
Met Thr Glu Phe Asn Tyr Asn Ser Val Met Gln Val Pro Arg Val Glu

```
            20                  25                  30
Lys Ile Thr Leu Asn Met Gly Val Gly Glu Ala Ile Ala Asp Lys Lys
        35                  40                  45
Leu Leu Asp Asn Ala Ala Ala Asp Leu Ala Ala Ile Ser Gly Gln Lys
    50                  55                  60
Pro Leu Ile Thr Lys Ala Arg Lys Ser Val Ala Gly Phe Lys Ile Arg
65                  70                  75                  80
Gln Gly Tyr Pro Ile Gly Cys Lys Val Thr Leu Arg Gly Glu Arg Met
                85                  90                  95
Trp Glu Phe Phe Glu Arg Leu Ile Thr Ile Ala Val Pro Arg Ile Arg
            100                 105                 110
Asp Phe Arg Gly Leu Ser Ala Lys Ser Phe Asp Gly Arg Gly Asn Tyr
        115                 120                 125
Ser Met Gly Val Arg Glu Gln Ile Ile Phe Pro Glu Ile Asp Tyr Asp
    130                 135                 140
Lys Val Asp Arg Val Arg Gly Leu Asp Ile Thr Ile Thr Thr Thr Ala
145                 150                 155                 160
Lys Ser Asp Glu Glu Gly Arg Ala Leu Leu Ala Ala Phe Asp Phe Pro
                165                 170                 175
Phe Arg Lys
```

<210> 297
<211> 104
<212> PRT
<213> Escherichia coli

<400> 297

```
Met Ala Ala Lys Ile Arg Arg Asp Asp Glu Val Ile Val Leu Thr Gly
 1               5                  10                  15
Lys Asp Lys Gly Lys Arg Gly Lys Val Lys Asn Val Leu Ser Ser Gly
            20                  25                  30
Lys Val Ile Val Glu Gly Ile Asn Leu Val Lys Lys His Gln Lys Pro
        35                  40                  45
Val Pro Ala Leu Asn Gln Pro Gly Gly Ile Val Glu Lys Glu Ala Ala
    50                  55                  60
Ile Gln Val Ser Asn Val Ala Ile Phe Asn Ala Ala Thr Gly Lys Ala
65                  70                  75                  80
Asp Arg Val Gly Phe Arg Phe Glu Asp Gly Lys Lys Val Arg Phe Phe
                85                  90                  95
Lys Ser Asn Ser Glu Thr Ile Lys
            100
```

<210> 298
<211> 123
<212> PRT
<213> Escherichia coli

<400> 298

```
Met Ile Gln Glu Gln Thr Met Leu Asn Val Ala Asp Asn Ser Gly Ala
 1               5                  10                  15
Arg Arg Val Met Cys Ile Lys Val Leu Gly Gly Ser His Arg Arg Tyr
            20                  25                  30
Ala Gly Val Gly Asp Ile Ile Lys Ile Thr Ile Lys Glu Ala Ile Pro
        35                  40                  45
Arg Gly Lys Val Lys Lys Gly Asp Val Leu Lys Ala Val Val Val Arg
    50                  55                  60
Thr Lys Lys Gly Val Arg Arg Pro Asp Gly Ser Val Ile Arg Phe Asp
65                  70                  75                  80
```

```
Gly Asn Ala Cys Val Leu Leu Asn Asn Asn Ser Glu Gln Pro Ile Gly
                85                  90                  95
Thr Arg Ile Phe Gly Pro Val Thr Arg Glu Leu Arg Ser Glu Lys Phe
            100                 105                 110
Met Lys Ile Ile Ser Leu Ala Pro Glu Val Leu
        115                 120


<210> 299
<211> 485
<212> PRT
<213> Escherichia coli

<400> 299
Met Gly Ile Tyr Phe Thr Asn Ser Asp Asp Gln Ile Tyr Phe Lys Arg
 1               5                  10                  15
Ser Glu Gly Met Ser Asp Ile Asn His Ala Gly Ser Asp Leu Ile Phe
            20                  25                  30
Glu Leu Glu Asp Arg Pro Pro Phe His Gln Ala Leu Val Gly Ala Ile
        35                  40                  45
Thr His Leu Leu Ala Ile Phe Val Pro Met Val Thr Pro Ala Leu Ile
    50                  55                  60
Val Gly Ala Ala Leu Gln Leu Ser Ala Glu Thr Thr Ala Tyr Leu Val
65                  70                  75                  80
Ser Met Ala Met Ile Ala Ser Gly Ile Gly Thr Trp Leu Gln Val Asn
                85                  90                  95
Arg Tyr Gly Ile Val Gly Ser Gly Leu Leu Ser Ile Gln Ser Val Asn
            100                 105                 110
Phe Ser Phe Val Thr Val Met Ile Ala Leu Gly Ser Ser Met Lys Ser
        115                 120                 125
Asp Gly Phe His Glu Glu Leu Ile Met Ser Ser Leu Leu Gly Val Ser
    130                 135                 140
Phe Val Gly Ala Phe Leu Val Val Gly Ser Ser Phe Ile Leu Pro Tyr
145                 150                 155                 160
Leu Arg Arg Val Ile Thr Pro Thr Val Ser Gly Ile Val Val Leu Met
                165                 170                 175
Ile Gly Leu Ser Leu Ile Lys Val Gly Ile Ile Asp Phe Gly Gly Gly
            180                 185                 190
Phe Ala Ala Lys Ser Ser Gly Thr Phe Gly Asn Tyr Glu His Leu Gly
        195                 200                 205
Val Gly Leu Leu Val Leu Ile Val Val Ile Gly Phe Asn Cys Cys Arg
    210                 215                 220
Ser Pro Leu Leu Arg Met Gly Gly Ile Ala Ile Gly Leu Cys Val Gly
225                 230                 235                 240
Tyr Ile Ala Ser Leu Cys Leu Gly Met Val Asp Phe Ser Ser Met Arg
                245                 250                 255
Asn Leu Pro Leu Ile Thr Ile Pro His Pro Phe Lys Tyr Gly Phe Ser
            260                 265                 270
Phe Ser Phe His Gln Phe Leu Val Val Gly Thr Ile Tyr Leu Leu Ser
        275                 280                 285
Val Leu Glu Ala Val Gly Asp Ile Thr Ala Thr Ala Met Val Ser Arg
    290                 295                 300
Arg Pro Ile Gln Gly Glu Glu Tyr Gln Ser Arg Leu Lys Gly Gly Val
305                 310                 315                 320
Leu Ala Asp Gly Leu Val Ser Val Ile Ala Ser Ala Val Gly Ser Leu
                325                 330                 335
Pro Leu Thr Thr Phe Ala Gln Asn Asn Gly Val Ile Gln Met Thr Gly
            340                 345                 350
Val Ala Ser Arg Tyr Val Gly Arg Thr Ile Ala Val Met Leu Val Ile
        355                 360                 365
Leu Gly Leu Phe Pro Met Ile Gly Gly Phe Phe Thr Thr Ile Pro Ser
```

370 375 380
Ala Val Leu Gly Gly Ala Met Thr Leu Met Phe Ser Met Ile Ala Ile
385 390 395 400
Ala Gly Ile Arg Ile Ile Ile Thr Asn Gly Leu Lys Arg Arg Glu Thr
405 410 415
Leu Ile Val Ala Thr Ser Leu Gly Leu Gly Leu Gly Val Ser Tyr Asp
420 425 430
Pro Glu Ile Phe Lys Ile Leu Pro Ala Ser Ile Tyr Val Leu Val Glu
435 440 445
Asn Pro Ile Cys Ala Gly Gly Leu Thr Ala Ile Leu Leu Asn Ile Ile
450 455 460
Leu Pro Gly Gly Tyr Arg Gln Glu Asn Val Leu Pro Gly Ile Thr Ser
465 470 475 480
Ala Glu Glu Met Asp
485

<210> 300
<211> 439
<212> PRT
<213> Escherichia coli

<400> 300
Met Met Ser Gly Glu His Thr Leu Lys Ala Val Arg Gly Ser Phe Ile
1 5 10 15
Asp Val Thr Arg Thr Ile Asp Asn Pro Glu Glu Ile Ala Ser Ala Leu
20 25 30
Arg Phe Ile Glu Asp Gly Leu Leu Leu Ile Lys Gln Gly Lys Val Glu
35 40 45
Trp Phe Gly Glu Trp Glu Asn Gly Lys His Gln Ile Pro Asp Thr Ile
50 55 60
Arg Val Arg Asp Tyr Arg Gly Lys Leu Ile Val Pro Gly Phe Val Asp
65 70 75 80
Thr His Ile His Tyr Pro Gln Ser Glu Met Val Gly Ala Tyr Gly Glu
85 90 95
Gln Leu Leu Glu Trp Leu Asn Lys His Thr Phe Pro Thr Glu Arg Arg
100 105 110
Tyr Glu Asp Leu Glu Tyr Ala Arg Glu Met Ser Ala Phe Phe Ile Lys
115 120 125
Gln Leu Leu Arg Asn Gly Thr Thr Thr Ala Leu Val Phe Gly Thr Val
130 135 140
His Pro Gln Ser Val Asp Ala Leu Phe Glu Ala Ala Ser His Ile Asn
145 150 155 160
Met Arg Met Ile Ala Gly Lys Val Met Met Asp Arg Asn Ala Pro Asp
165 170 175
Tyr Leu Leu Asp Thr Ala Glu Ser Ser Tyr His Gln Ser Lys Glu Leu
180 185 190
Ile Glu Arg Trp His Lys Asn Gly Arg Leu Leu Tyr Ala Ile Thr Pro
195 200 205
Arg Phe Ala Pro Thr Ser Ser Pro Glu Gln Met Ala Met Ala Gln Arg
210 215 220
Leu Lys Glu Glu Tyr Pro Asp Thr Trp Val His Thr His Leu Cys Glu
225 230 235 240
Asn Lys Asp Glu Ile Ala Trp Val Lys Ser Leu Tyr Pro Asp His Asp
245 250 255
Gly Tyr Leu Asp Val Tyr His Gln Tyr Gly Leu Thr Gly Lys Asn Cys
260 265 270
Val Phe Ala His Cys Val His Leu Glu Glu Lys Glu Trp Asp Arg Leu
275 280 285
Ser Glu Thr Lys Ser Ser Ile Ala Phe Cys Pro Thr Ser Asn Leu Tyr
290 295 300

Leu Gly Ser Gly Leu Phe Asn Leu Lys Lys Ala Trp Gln Lys Lys Val
305 310 315 320
Lys Val Gly Met Gly Thr Asp Ile Gly Ala Gly Thr Thr Phe Asn Met
325 330 335
Leu Gln Thr Leu Asn Glu Ala Tyr Lys Val Leu Gln Leu Gln Gly Tyr
340 345 350
Arg Leu Ser Ala Tyr Glu Ala Phe Tyr Leu Ala Thr Leu Gly Gly Ala
355 360 365
Lys Ser Leu Gly Leu Asp Asp Leu Ile Gly Asn Phe Leu Pro Gly Lys
370 375 380
Glu Ala Asp Phe Val Val Met Glu Pro Thr Ala Thr Pro Leu Gln Gln
385 390 395 400
Leu Arg Tyr Asp Asn Ser Val Ser Leu Val Asp Lys Leu Phe Val Met
405 410 415
Met Thr Leu Gly Asp Asp Arg Ser Ile Tyr Arg Thr Tyr Val Asp Gly
420 425 430
Arg Leu Val Tyr Glu Arg Asn
435

<210> 301
<211> 189
<212> PRT
<213> Escherichia coli

<400> 301
Met Ser Gly Asp Ile Leu Gln Thr Pro Asp Ala Pro Lys Pro Gln Gly
1 5 10 15
Ala Leu Asp Asn Tyr Phe Lys Ile Thr Ala Arg Gly Ser Thr Val Arg
20 25 30
Gln Glu Val Leu Ala Gly Leu Thr Thr Phe Leu Ala Met Val Tyr Ser
35 40 45
Val Ile Val Val Pro Gly Met Leu Gly Lys Ala Gly Phe Pro Pro Ala
50 55 60
Ala Val Phe Val Ala Thr Cys Leu Val Ala Gly Phe Gly Ser Leu Leu
65 70 75 80
Met Gly Leu Trp Ala Asn Leu Pro Met Ala Ile Gly Cys Ala Ile Ser
85 90 95
Leu Thr Ala Phe Thr Ala Phe Ser Leu Val Leu Gly Gln Gln Ile Ser
100 105 110
Val Pro Val Ala Leu Gly Ala Val Phe Leu Met Gly Val Ile Phe Thr
115 120 125
Ala Ile Ser Val Thr Gly Val Arg Thr Trp Ile Leu Arg Asn Leu Pro
130 135 140
Met Gly Ile Ala His Gly Thr Gly Ile Gly Ile Gly Leu Phe Leu Leu
145 150 155 160
Leu Ile Ala Ala Asn Gly Val Gly Met Val Ile Lys Asn Pro Ile Glu
165 170 175
Gly Leu Gln Trp Arg Ser Val Arg Leu Pro Pro Ser Arg
180 185

<210> 302
<211> 276
<212> PRT
<213> Escherichia coli

<400> 302
Met Ala Leu Gly Ala Phe Thr Ser Phe Pro Val Met Met Ser Leu Leu
1 5 10 15
Gly Leu Ala Val Ile Phe Gly Leu Glu Lys Cys Arg Val Pro Gly Gly

```
            20                  25                  30
Ile Leu Leu Val Ile Ile Ala Ile Ser Ile Ile Gly Leu Ile Phe Asp
        35                  40                  45
Pro Ala Val Lys Tyr His Gly Leu Val Ala Met Pro Ser Leu Thr Gly
    50                  55                  60
Glu Asp Gly Lys Ser Leu Ile Phe Ser Leu Asp Ile Met Gly Ala Leu
65                  70                  75                  80
Gln Pro Thr Val Leu Pro Ser Val Leu Ala Leu Val Met Thr Ala Val
                85                  90                  95
Phe Asp Ala Thr Gly Thr Ile Arg Ala Val Ala Gly Gln Ala Asn Leu
            100                 105                 110
Leu Asp Lys Asp Asn Gln Ile Ile Asn Gly Gly Lys Ala Leu Thr Ser
        115                 120                 125
Asp Ser Val Ser Ser Ile Phe Ser Gly Leu Val Gly Ala Ala Pro Ala
    130                 135                 140
Ala Val Tyr Ile Glu Ser Ala Ala Gly Thr Ala Ala Gly Gly Lys Thr
145                 150                 155                 160
Gly Leu Thr Ala Thr Val Val Gly Ala Leu Phe Leu Leu Ile Leu Phe
                165                 170                 175
Leu Ser Pro Leu Ser Phe Leu Ile Pro Gly Tyr Ala Thr Ala Pro Ala
            180                 185                 190
Leu Met Tyr Val Gly Leu Leu Met Leu Ser Asn Val Ser Lys Leu Asp
        195                 200                 205
Phe Asn Asp Phe Ile Asp Ala Met Ala Gly Leu Val Cys Ala Val Phe
    210                 215                 220
Ile Val Leu Thr Cys Asn Ile Val Thr Gly Ile Met Leu Gly Phe Val
225                 230                 235                 240
Thr Leu Val Val Gly Arg Val Phe Ala Arg Glu Trp Gln Lys Leu Asn
                245                 250                 255
Ile Gly Thr Val Ile Ile Thr Ala Ala Leu Val Ala Phe Tyr Ala Gly
            260                 265                 270
Gly Trp Ala Ile
        275
```

<210> 303
<211> 466
<212> PRT
<213> Escherichia coli

<400> 303

```
Met Asn Ser Glu Gly Gly Lys Pro Gly Asn Val Leu Thr Val Asn Gly
 1               5                  10                  15
Asn Tyr Thr Gly Asn Asn Gly Leu Met Thr Phe Asn Ala Thr Leu Gly
            20                  25                  30
Gly Asp Asn Ser Pro Thr Asp Lys Met Asn Val Lys Gly Asp Thr Gln
        35                  40                  45
Gly Asn Thr Arg Val Arg Val Asp Asn Ile Gly Gly Val Gly Ala Gln
    50                  55                  60
Thr Val Asn Gly Ile Glu Leu Ile Glu Val Gly Gly Asn Ser Ala Gly
65                  70                  75                  80
Asn Phe Ala Leu Thr Thr Gly Thr Val Glu Ala Gly Ala Tyr Val Tyr
                85                  90                  95
Thr Leu Ala Lys Gly Lys Gly Asn Asp Glu Lys Asn Trp Tyr Leu Thr
            100                 105                 110
Ser Lys Trp Asp Gly Val Thr Pro Ala Asp Thr Pro Asp Pro Ile Asn
        115                 120                 125
Asn Pro Pro Val Val Asp Pro Glu Gly Pro Ser Val Tyr Arg Pro Glu
    130                 135                 140
Ala Gly Ser Tyr Ile Ser Asn Ile Ala Ala Ala Asn Ser Leu Phe Ser
145                 150                 155                 160
```

His Arg Leu His Asp Arg Leu Gly Glu Pro Gln Tyr Thr Asp Ser Leu
165 170 175
His Ser Gln Gly Ser Ala Ser Ser Met Trp Met Arg His Val Gly Gly
180 185 190
His Glu Arg Ser Arg Ala Gly Asp Gly Gln Leu Asn Thr Gln Ala Asn
195 200 205
Arg Tyr Val Leu Gln Leu Gly Gly Asp Leu Ala Gln Trp Ser Ser Asn
210 215 220
Ala Gln Asp Arg Trp His Leu Gly Val Met Ala Gly Tyr Ala Asn Gln
225 230 235 240
His Ser Asn Thr Gln Ser Asn Arg Val Gly Tyr Lys Ser Asp Gly Arg
245 250 255
Ile Ser Gly Tyr Ser Ala Gly Leu Tyr Ala Thr Trp Tyr Gln Asn Asp
260 265 270
Ala Asn Lys Thr Gly Ala Tyr Val Asp Ser Trp Ala Leu Tyr Asn Trp
275 280 285
Phe Asp Asn Ser Val Ser Ser Asp Asn Arg Ser Ala Asp Asp Tyr Asp
290 295 300
Ser Arg Gly Val Thr Ala Ser Val Glu Gly Gly Tyr Thr Phe Glu Ala
305 310 315 320
Gly Thr Phe Ser Gly Ser Glu Gly Thr Leu Asn Thr Trp Tyr Val Gln
325 330 335
Pro Gln Ala Gln Ile Thr Trp Met Gly Val Lys Asp Ser Asp His Thr
340 345 350
Arg Lys Asp Gly Thr Arg Ile Glu Thr Glu Gly Asp Gly Asn Val Gln
355 360 365
Thr Arg Leu Gly Val Lys Thr Tyr Leu Asn Ser His His Gln Arg Asp
370 375 380
Asp Gly Lys Gln Arg Glu Phe Gln Pro Tyr Ile Glu Ala Asn Trp Ile
385 390 395 400
Asn Asn Ser Lys Val Tyr Ala Val Lys Met Asn Gly Gln Thr Val Gly
405 410 415
Arg Glu Gly Ala Arg Asn Leu Gly Glu Val Arg Thr Gly Val Glu Ala
420 425 430
Lys Val Asn Asn Asn Leu Ser Leu Trp Gly Asn Val Gly Val Gln Leu
435 440 445
Gly Asp Lys Gly Tyr Ser Asp Thr Gln Gly Met Leu Gly Val Lys Tyr
450 455 460
Ser Trp
465

<210> 304
<211> 1325
<212> PRT
<213> Escherichia coli

<400> 304
Met Asn Arg Ile Tyr Arg Val Ile Trp Asn Cys Thr Leu Gln Val Phe
1 5 10 15
Gln Ala Cys Ser Glu Leu Thr Arg Arg Ala Gly Lys Thr Ser Thr Val
20 25 30
Asn Leu Arg Lys Ser Ser Gly Leu Thr Thr Lys Phe Ser Arg Leu Thr
35 40 45
Leu Gly Val Leu Leu Ala Leu Ser Gly Ser Ala Ser Gly Ala Ser Leu
50 55 60
Glu Val Asp Asn Asp Gln Ile Thr Asn Ile Asp Thr Asp Val Ala Tyr
65 70 75 80
Asp Ala Tyr Leu Val Gly Trp Tyr Gly Thr Gly Val Leu Asn Ile Leu
85 90 95
Ala Gly Gly Asn Ala Ser Leu Thr Thr Ile Thr Thr Ser Val Ile Gly

```
            100                     105                     110
Ala Asn Glu Asp Ser Glu Gly Thr Val Asn Val Leu Gly Gly Thr Trp
        115                     120                     125
Arg Leu Tyr Asp Ser Gly Asn Asn Ala Arg Pro Leu Asn Val Gly Gln
    130                     135                     140
Ser Gly Thr Gly Thr Leu Asn Ile Lys Gln Lys Gly His Val Asp Gly
145                     150                     155                     160
Gly Tyr Leu Arg Leu Gly Ser Ser Thr Gly Gly Val Gly Thr Val Asn
                165                     170                     175
Val Glu Gly Glu Asp Ser Val Leu Thr Thr Glu Leu Phe Glu Ile Gly
            180                     185                     190
Ser Tyr Gly Thr Gly Ser Leu Asn Ile Thr Asp Lys Gly Tyr Val Thr
        195                     200                     205
Ser Ser Ile Val Ala Ile Leu Gly Tyr Gln Ala Gly Ser Asn Gly Gln
    210                     215                     220
Val Val Val Glu Lys Gly Gly Glu Trp Leu Ile Lys Asn Asn Asp Ser
225                     230                     235                     240
Ser Ile Glu Phe Gln Ile Gly Asn Gln Gly Thr Gly Glu Ala Thr Ile
                245                     250                     255
Arg Glu Gly Gly Leu Val Thr Ala Glu Asn Thr Ile Ile Gly Gly Asn
            260                     265                     270
Ala Thr Gly Ile Gly Thr Leu Asn Val Gln Asp Gln Asp Ser Val Ile
        275                     280                     285
Thr Val Arg Arg Leu Tyr Asn Gly Tyr Phe Gly Asn Gly Thr Val Asn
    290                     295                     300
Ile Ser Asn Asn Gly Leu Ile Asn Asn Lys Glu Tyr Ser Leu Val Gly
305                     310                     315                     320
Val Gln Asp Gly Ser His Gly Val Val Asn Val Thr Asp Lys Gly His
                325                     330                     335
Trp Asn Phe Leu Gly Thr Gly Glu Ala Phe Arg Tyr Ile Tyr Ile Gly
            340                     345                     350
Asp Ala Gly Asp Gly Glu Leu Asn Val Ser Ser Glu Gly Lys Val Asp
        355                     360                     365
Ser Gly Ile Ile Thr Ala Gly Met Lys Glu Thr Gly Thr Gly Asn Ile
    370                     375                     380
Thr Val Lys Asp Lys Asn Ser Val Ile Thr Asn Leu Gly Thr Asn Leu
385                     390                     395                     400
Gly Tyr Asp Gly His Gly Glu Met Asn Ile Ser Asn Gln Gly Leu Val
                405                     410                     415
Val Ser Asn Gly Gly Ser Ser Leu Gly Tyr Gly Glu Thr Gly Val Gly
            420                     425                     430
Asn Val Ser Ile Thr Thr Gly Gly Met Trp Glu Val Asn Lys Asn Val
        435                     440                     445
Tyr Thr Thr Ile Gly Val Ala Gly Val Gly Asn Leu Asn Ile Ser Asp
    450                     455                     460
Gly Gly Lys Phe Val Ser Gln Asn Ile Thr Phe Leu Gly Asp Lys Ala
465                     470                     475                     480
Ser Gly Ile Gly Thr Leu Asn Leu Met Asp Ala Thr Ser Ser Phe Asp
                485                     490                     495
Thr Val Gly Ile Asn Val Gly Asn Phe Gly Ser Gly Ile Val Asn Val
            500                     505                     510
Ser Asn Gly Ala Thr Leu Asn Ser Thr Gly Tyr Gly Phe Ile Gly Gly
        515                     520                     525
Asn Ala Ser Gly Lys Gly Ile Val Asn Ile Ser Thr Asp Ser Leu Trp
    530                     535                     540
Asn Leu Lys Thr Ser Ser Thr Asn Ala Gln Leu Leu Gln Val Gly Val
545                     550                     555                     560
Leu Gly Thr Gly Glu Leu Asn Ile Thr Thr Gly Gly Ile Val Lys Ala
                565                     570                     575
Arg Asp Thr Gln Ile Ala Leu Asn Asp Lys Ser Lys Gly Asp Val Arg
            580                     585                     590
```

Val Asp Gly Gln Asn Ser Leu Leu Glu Thr Phe Asn Met Tyr Val Gly
595 600 605
Thr Ser Gly Thr Gly Thr Leu Thr Leu Thr Asn Asn Gly Thr Leu Asn
610 615 620
Val Glu Gly Gly Glu Val Tyr Leu Gly Val Phe Glu Pro Ala Val Gly
625 630 635 640
Thr Leu Asn Ile Gly Ala Ala His Gly Glu Ala Ala Ala Asp Ala Gly
645 650 655
Phe Ile Thr Asn Ala Thr Lys Val Glu Phe Gly Leu Gly Glu Gly Val
660 665 670
Phe Val Phe Asn His Thr Asn Asn Ser Asp Ala Gly Tyr Gln Val Asp
675 680 685
Met Leu Ile Thr Gly Asp Asp Lys Asp Gly Lys Val Ile His Asp Ala
690 695 700
Gly His Thr Val Phe Asn Ala Gly Asn Thr Tyr Ser Gly Lys Thr Leu
705 710 715 720
Val Asn Asp Gly Leu Leu Thr Ile Ala Ser His Thr Ala Asp Gly Val
725 730 735
Thr Gly Met Gly Ser Ser Glu Val Thr Ile Ala Asn Pro Gly Thr Leu
740 745 750
Asp Ile Leu Ala Ser Thr Asn Ser Ala Gly Asp Tyr Thr Leu Thr Asn
755 760 765
Ala Leu Lys Gly Asp Gly Leu Met Arg Val Gln Leu Ser Ser Ser Asp
770 775 780
Lys Met Phe Gly Phe Thr His Ala Thr Gly Thr Glu Phe Ala Gly Val
785 790 795 800
Ala Gln Leu Lys Asp Ser Thr Phe Thr Leu Glu Arg Asp Asn Thr Ala
805 810 815
Ala Leu Thr His Ala Met Leu Gln Ser Asp Ser Glu Asn Thr Thr Ser
820 825 830
Val Lys Val Gly Glu Gln Ser Ile Gly Gly Leu Ala Met Asn Gly Gly
835 840 845
Thr Ile Ile Phe Asp Thr Asp Ile Pro Ala Ala Thr Leu Ala Glu Gly
850 855 860
Tyr Ile Ser Val Asp Thr Leu Val Val Gly Ala Gly Asp Tyr Thr Trp
865 870 875 880
Lys Gly Arg Asn Tyr Gln Val Asn Gly Thr Gly Asp Val Leu Ile Asp
885 890 895
Val Pro Lys Pro Trp Asn Asp Pro Met Ala Asn Asn Pro Leu Thr Thr
900 905 910
Leu Asn Leu Leu Glu His Asp Asp Ser His Val Gly Val Gln Leu Val
915 920 925
Lys Ala Gln Thr Val Ile Gly Ser Gly Gly Ser Leu Thr Leu Arg Asp
930 935 940
Leu Gln Gly Asp Glu Val Glu Ala Asp Lys Thr Leu His Ile Ala Gln
945 950 955 960
Asn Gly Thr Val Val Ala Glu Gly Asp Tyr Gly Phe Arg Leu Thr Thr
965 970 975
Ala Pro Gly Asn Gly Leu Tyr Val Asn Tyr Gly Leu Lys Ala Leu Asn
980 985 990
Ile His Gly Gly Gln Lys Leu Thr Leu Ala Glu His Gly Gly Ala Tyr
995 1000 1005
Gly Ala Thr Ala Asp Met Ser Ala Lys Ile Gly Gly Glu Gly Asp Leu
1010 1015 1020
Ala Ile Asn Thr Val Arg Gln Val Ser Leu Ser Asn Gly Gln Asn Asp
1025 1030 1035 1040
Tyr Gln Gly Ala Thr Tyr Val Gln Met Gly Thr Leu Arg Thr Asp Ala
1045 1050 1055
Asp Gly Ala Leu Gly Asn Thr Arg Glu Leu Asn Ile Ser Asn Ala Ala
1060 1065 1070
Ile Val Asp Leu Asn Gly Ser Thr Gln Thr Val Glu Thr Phe Thr Gly

1075 1080 1085
Gln Met Gly Ser Thr Val Leu Phe Lys Glu Gly Ala Leu Thr Val Asn
1090 1095 1100
Lys Gly Gly Ile Ser Gln Gly Glu Leu Thr Gly Gly Gly Asn Leu Asn
1105 1110 1115 1120
Val Thr Gly Gly Thr Leu Ala Ile Glu Gly Leu Asn Ala Arg Tyr Asn
1125 1130 1135
Ala Leu Thr Ser Ile Ser Pro Asn Ala Glu Val Ser Leu Asp Asn Thr
1140 1145 1150
Gln Gly Leu Gly Arg Gly Asn Ile Ala Asn Asp Gly Leu Leu Thr Leu
1155 1160 1165
Lys Asn Val Thr Gly Glu Leu Arg Asn Ser Ile Ser Gly Lys Gly Ile
1170 1175 1180
Val Ser Ala Thr Ala Arg Thr Asp Val Glu Leu Asp Gly Asp Asn Ser
1185 1190 1195 1200
Arg Phe Val Gly Gln Phe Asn Ile Asp Thr Gly Ser Ala Leu Ser Val
1205 1210 1215
Asn Glu Gln Lys Asn Leu Gly Asp Ala Ser Val Ile Asn Asn Gly Leu
1220 1225 1230
Leu Thr Ile Ser Thr Glu Arg Ser Trp Ala Met Thr His Ser Ile Ser
1235 1240 1245
Gly Ser Gly Asp Val Thr Lys Leu Gly Thr Gly Ile Leu Thr Leu Asn
1250 1255 1260
Asn Asp Ser Ala Ala Tyr Gln Gly Thr Thr Asp Ile Val Gly Gly Glu
1265 1270 1275 1280
Ile Ala Phe Gly Ser Asp Ser Ala Ile Asn Met Ala Ser Gln His Ile
1285 1290 1295
Asn Ile His Asn Ser Gly Val Met Ser Gly Asn Val Thr Thr Ala Gly
1300 1305 1310
Asp Met Asn Val Met Pro Gly Gly Gly Thr Ala Cys Arg
1315 1320 1325

<210> 305
<211> 251
<212> PRT
<213> Escherichia coli

<400> 305
Met Thr Glu Ala Gln Arg His Gln Ile Leu Leu Glu Met Leu Ala Gln
1 5 10 15
Leu Gly Phe Val Thr Val Glu Lys Val Val Glu Arg Leu Gly Ile Ser
20 25 30
Pro Ala Thr Ala Arg Arg Asp Ile Asn Lys Leu Asp Glu Ser Gly Lys
35 40 45
Leu Lys Lys Val Arg Asn Gly Ala Glu Ala Ile Thr Gln Gln Arg Pro
50 55 60
Arg Trp Thr Pro Met Asn Leu His Gln Ala Gln Asn His Asp Glu Lys
65 70 75 80
Val Arg Ile Ala Lys Ala Ala Ser Gln Leu Val Asn Pro Gly Glu Ser
85 90 95
Val Val Ile Asn Cys Gly Ser Thr Ala Phe Leu Leu Gly Arg Glu Met
100 105 110
Cys Gly Lys Pro Val Gln Ile Ile Thr Asn Tyr Leu Pro Leu Ala Asn
115 120 125
Tyr Leu Ile Asp Gln Glu His Asp Ser Val Ile Ile Met Gly Gly Gln
130 135 140
Tyr Asn Lys Ser Gln Ser Ile Thr Leu Ser Pro Gln Gly Ser Glu Asn
145 150 155 160
Ser Leu Tyr Ala Gly His Trp Met Phe Thr Ser Gly Lys Gly Leu Thr
165 170 175

Ala Glu Gly Leu Tyr Lys Thr Asp Met Leu Thr Ala Met Ala Glu Gln
180 185 190
Lys Met Leu Ser Val Val Gly Lys Leu Val Val Leu Val Asp Ser Ser
195 200 205
Lys Ile Gly Glu Arg Ala Gly Met Leu Phe Ser Arg Ala Asp Gln Ile
210 215 220
Asp Met Leu Ile Thr Gly Lys Asn Ala Asn Pro Glu Ile Leu Gln Gln
225 230 235 240
Leu Glu Ala Gln Gly Val Ser Ile Leu Arg Val
245 250

<210> 306
<211> 274
<212> PRT
<213> Escherichia coli

<400> 306
Met Thr Glu Phe Thr Thr Leu Leu Gln Gln Gly Asn Ala Trp Phe Phe
1 5 10 15
Ile Pro Ser Ala Ile Leu Leu Gly Ala Leu His Gly Leu Glu Pro Gly
20 25 30
His Ser Lys Thr Met Met Ala Ala Phe Ile Ile Ala Ile Lys Gly Thr
35 40 45
Ile Lys Gln Ala Val Met Leu Gly Leu Ala Ala Thr Ile Ser His Thr
50 55 60
Ala Val Val Trp Leu Ile Ala Phe Gly Gly Met Val Ile Ser Lys Arg
65 70 75 80
Phe Thr Ala Gln Ser Ala Glu Pro Trp Leu Gln Leu Ile Ser Ala Val
85 90 95
Ile Ile Ile Ser Thr Ala Phe Trp Met Phe Trp Arg Thr Trp Arg Gly
100 105 110
Glu Arg Asn Trp Leu Glu Asn Met His Gly His Asp Tyr Glu His His
115 120 125
His His Asp His Glu His His His Asp His Gly His His His His His
130 135 140
Glu His Gly Glu Tyr Gln Asp Ala His Ala Arg Ala His Ala Asn Asp
145 150 155 160
Ile Lys Arg Arg Phe Asp Gly Arg Glu Val Thr Asn Trp Gln Ile Leu
165 170 175
Leu Phe Gly Leu Thr Gly Gly Leu Ile Pro Cys Pro Ala Ala Ile Thr
180 185 190
Val Leu Leu Ile Cys Ile Gln Leu Lys Ala Leu Thr Leu Gly Ala Thr
195 200 205
Leu Val Val Ser Phe Ser Ile Gly Leu Ala Leu Thr Leu Val Thr Val
210 215 220
Gly Val Gly Ala Ala Ile Ser Val Gln Gln Val Ala Lys Arg Trp Ser
225 230 235 240
Gly Phe Asn Thr Leu Ala Lys Arg Ala Pro Tyr Phe Ser Ser Leu Leu
245 250 255
Ile Gly Leu Val Gly Val Tyr Met Gly Val His Gly Phe Met Gly Ile
260 265 270
Met Arg

<210> 307
<211> 172
<212> PRT
<213> Escherichia coli

<400> 307
Met Ile Leu Lys Ser Ala Ile Ser Ala Asp Ser Leu Leu Ala Lys Asp
1 5 10 15
Ala Phe Arg Ala Ser Phe His Leu His Phe Leu Arg Asn His Gly Ile
20 25 30
Thr Asn Lys Ile Ser Leu Val Ser Tyr Ile Val Trp Gln Glu Arg Tyr
35 40 45
Ala Thr Asp Ile Thr Asp Pro Gln Ser Gly Glu Phe Met Thr Ile Lys
50 55 60
Asn Lys Met Leu Leu Gly Ala Leu Leu Leu Val Thr Ser Ala Ala Trp
65 70 75 80
Ala Ala Pro Ala Thr Ala Gly Ser Thr Asn Thr Ser Gly Ile Ser Lys
85 90 95
Tyr Glu Leu Ser Ser Phe Ile Ala Asp Phe Lys His Phe Lys Pro Gly
100 105 110
Asp Thr Val Pro Glu Met Tyr Arg Thr Asp Glu Tyr Asn Ile Lys Gln
115 120 125
Trp Gln Leu Arg Asn Leu Pro Ala Pro Asp Ala Gly Thr His Trp Thr
130 135 140
Tyr Met Gly Gly Ala Tyr Val Leu Ile Ser Asp Thr Asp Gly Lys Ile
145 150 155 160
Ile Lys Ala Tyr Asp Gly Glu Ile Phe Tyr His Arg
165 170

<210> 308
<211> 344
<212> PRT
<213> Escherichia coli

<400> 308
Met Glu Ile Arg Ile Met Leu Phe Ile Leu Met Met Met Val Met Pro
1 5 10 15
Val Ser Tyr Ala Ala Cys Tyr Ser Glu Leu Ser Val Gln His Asn Leu
20 25 30
Val Val Gln Gly Asp Phe Ala Leu Thr Gln Thr Gln Met Ala Thr Tyr
35 40 45
Glu His Asn Phe Asn Asp Ser Ser Cys Val Ser Thr Asn Thr Ile Thr
50 55 60
Pro Met Ser Pro Ser Asp Ile Ile Val Gly Leu Tyr Asn Asp Thr Ile
65 70 75 80
Lys Leu Asn Leu His Phe Glu Trp Thr Asn Lys Asn Asn Ile Thr Leu
85 90 95
Ser Asn Asn Gln Thr Ser Phe Thr Ser Gly Tyr Ser Val Thr Val Thr
100 105 110
Pro Ala Ala Ser Asn Ala Lys Val Asn Val Ser Ala Gly Gly Gly Gly
115 120 125
Ser Val Met Ile Asn Gly Val Ala Thr Leu Ser Ser Ala Ser Ser Ser
130 135 140
Thr Arg Gly Ser Ala Ala Val Gln Phe Leu Leu Cys Leu Leu Gly Gly
145 150 155 160
Lys Ser Trp Asp Ala Cys Val Asn Ser Tyr Arg Asn Ala Leu Ala Gln
165 170 175
Asn Ala Gly Val Tyr Ser Phe Asn Leu Thr Leu Ser Tyr Asn Pro Ile
180 185 190
Thr Thr Thr Cys Lys Pro Asp Asp Leu Leu Ile Thr Leu Asp Ser Ile
195 200 205
Pro Val Ser Gln Leu Pro Ala Thr Gly Asn Lys Ala Thr Ile Asn Ser
210 215 220
Lys Gln Gly Asp Ile Ile Leu Arg Cys Lys Asn Leu Leu Gly Gln Gln
225 230 235 240

Asn Gln Thr Ser Arg Lys Met Gln Val Tyr Leu Ser Ser Ser Asp Leu
245 250 255
Leu Thr Asn Ser Asn Thr Ile Leu Lys Gly Ala Glu Asp Asn Gly Val
260 265 270
Gly Phe Ile Leu Glu Ser Asn Gly Ser Pro Val Thr Leu Leu Asn Ile
275 280 285
Thr Asn Ser Ser Lys Gly Tyr Thr Asn Leu Lys Glu Val Ala Ala Lys
290 295 300
Ser Lys Leu Thr Asp Thr Thr Val Ser Ile Pro Ile Thr Ala Ser Tyr
305 310 315 320
Tyr Val Tyr Asp Thr Asn Lys Val Lys Ser Gly Ala Leu Glu Ala Thr
325 330 335
Ala Leu Ile Asn Val Lys Tyr Asp
340

<210> 309
<211> 826
<212> PRT
<213> Escherichia coli

<400> 309
Met Leu Arg Met Thr Pro Leu Ala Ser Ala Ile Val Ala Leu Leu Leu
1 5 10 15
Gly Ile Glu Ala Tyr Ala Ala Glu Glu Thr Phe Asp Thr His Phe Met
20 25 30
Ile Gly Gly Met Lys Asp Gln Gln Val Ala Asn Ile Arg Leu Asp Asp
35 40 45
Asn Gln Pro Leu Pro Gly Gln Tyr Asp Ile Asp Ile Tyr Val Asn Lys
50 55 60
Gln Trp Arg Gly Lys Tyr Glu Ile Ile Val Lys Asp Asn Pro Gln Glu
65 70 75 80
Thr Cys Leu Ser Arg Glu Val Ile Lys Arg Leu Gly Ile Asn Ser Asp
85 90 95
Asn Phe Ala Ser Gly Lys Gln Cys Leu Thr Phe Glu Gln Leu Val Gln
100 105 110
Gly Gly Ser Tyr Thr Trp Asp Ile Gly Val Phe Arg Leu Asp Phe Ser
115 120 125
Val Pro Gln Ala Trp Val Glu Glu Leu Glu Ser Gly Tyr Val Pro Pro
130 135 140
Glu Asn Trp Glu Arg Gly Ile Asn Ala Phe Tyr Thr Ser Tyr Tyr Leu
145 150 155 160
Ser Gln Tyr Tyr Ser Asp Tyr Lys Ala Ser Gly Asn Asn Lys Ser Thr
165 170 175
Tyr Val Arg Phe Asn Ser Gly Leu Asn Leu Leu Gly Trp Gln Leu His
180 185 190
Ser Asp Ala Ser Phe Ser Lys Thr Asn Asn Asn Pro Gly Val Trp Lys
195 200 205
Ser Asn Thr Leu Tyr Leu Glu Arg Gly Phe Ala Gln Leu Leu Gly Thr
210 215 220
Leu Arg Val Gly Asp Met Tyr Thr Ser Ser Asp Ile Phe Asp Ser Val
225 230 235 240
Arg Phe Arg Gly Val Arg Leu Phe Arg Asp Met Gln Met Leu Pro Asn
245 250 255
Ser Lys Gln Asn Phe Thr Pro Arg Val Gln Gly Ile Ala Gln Ser Asn
260 265 270
Ala Leu Val Thr Ile Glu Gln Asn Gly Phe Val Val Tyr Gln Lys Glu
275 280 285
Val Pro Pro Gly Pro Phe Ala Ile Thr Asp Leu Gln Leu Ala Gly Gly
290 295 300
Gly Ala Asp Leu Asp Val Ser Val Lys Glu Ala Asp Gly Ser Val Thr

305 310 315 320
Thr Tyr Leu Val Pro Tyr Ala Ala Val Pro Asn Met Leu Gln Pro Gly
325 330 335
Val Ser Lys Tyr Asp Leu Ala Ala Gly Arg Ser His Ile Glu Gly Ala
340 345 350
Ser Lys Gln Ser Asp Phe Val Gln Ala Gly Tyr Gln Tyr Gly Phe Asn
355 360 365
Asn Leu Leu Thr Leu Tyr Gly Gly Ser Met Val Ala Asn Asn Tyr Tyr
370 375 380
Ala Phe Thr Leu Gly Ala Gly Trp Asn Thr Arg Ile Gly Ala Ile Ser
385 390 395 400
Val Asp Ala Thr Lys Ser His Ser Lys Gln Asp Asn Gly Asp Val Phe
405 410 415
Asp Gly Gln Ser Tyr Gln Ile Ala Tyr Asn Lys Phe Val Ser Gln Thr
420 425 430
Ser Thr Arg Phe Gly Leu Ala Ala Trp Arg Tyr Ser Ser Arg Asp Tyr
435 440 445
Arg Thr Phe Asn Asp His Val Trp Ala Asn Asn Lys Asp Asn Tyr Arg
450 455 460
Arg Asp Glu Asn Asp Val Tyr Asp Ile Ala Asp Tyr Tyr Gln Asn Asp
465 470 475 480
Phe Gly Arg Lys Asn Ser Phe Ser Ala Asn Met Ser Gln Ser Leu Pro
485 490 495
Glu Gly Trp Gly Ser Val Ser Leu Ser Thr Leu Trp Arg Asp Tyr Trp
500 505 510
Gly Arg Ser Gly Ser Ser Lys Asp Tyr Gln Leu Ser Tyr Ser Asn Asn
515 520 525
Leu Arg Arg Ile Ser Tyr Thr Leu Ala Ala Ser Gln Ala Tyr Asp Glu
530 535 540
Asn His His Glu Glu Lys Arg Phe Asn Ile Phe Ile Ser Ile Pro Phe
545 550 555 560
Asp Trp Gly Asp Asp Val Ser Thr Pro Arg Arg Gln Ile Tyr Met Ser
565 570 575
Asn Ser Thr Thr Phe Asp Asp Gln Gly Phe Ala Ser Asn Asn Thr Gly
580 585 590
Leu Ser Gly Thr Val Gly Ser Arg Asp Gln Phe Asn Tyr Gly Val Asn
595 600 605
Leu Ser His Gln His Gln Gly Asn Glu Thr Thr Ala Gly Ala Asn Leu
610 615 620
Thr Trp Asn Ala Pro Val Ala Thr Val Asn Gly Ser Tyr Ser Gln Ser
625 630 635 640
Ser Thr Tyr Arg Gln Ala Gly Ala Ser Val Ser Gly Gly Ile Val Ala
645 650 655
Trp Ser Gly Gly Val Asn Leu Ala Asn Arg Leu Ser Glu Thr Phe Ala
660 665 670
Val Met Asn Ala Pro Gly Ile Lys Asp Ala Tyr Val Asn Gly Gln Lys
675 680 685
Tyr Arg Thr Thr Asn Arg Asn Gly Val Val Ile Tyr Asp Gly Met Thr
690 695 700
Pro Tyr Arg Glu Asn His Leu Met Leu Asp Val Ser Gln Ser Asp Ser
705 710 715 720
Glu Ala Glu Leu Arg Gly Asn Arg Lys Ile Ala Ala Pro Tyr Arg Gly
725 730 735
Ala Val Val Leu Val Asn Phe Asp Thr Asp Gln Arg Lys Pro Trp Phe
740 745 750
Ile Lys Ala Leu Arg Ala Asp Gly Gln Ser Leu Thr Phe Gly Tyr Glu
755 760 765
Val Asn Asp Ile His Gly His Asn Ile Gly Val Val Gly Gln Gly Ser
770 775 780
Gln Leu Phe Ile Arg Thr Asn Glu Val Pro Pro Ser Val Asn Val Ala
785 790 795 800

Ile Asp Lys Gln Gln Gly Leu Ser Cys Thr Ile Thr Phe Gly Lys Glu
805 810 815
Ile Asp Glu Ser Arg Asn Tyr Ile Cys Gln
820 825

<210> 310
<211> 239
<212> PRT
<213> Escherichia coli

<400> 310
Met Ala Ala Ile Pro Trp Arg Pro Phe Asn Leu Arg Gly Ile Lys Met
1 5 10 15
Lys Gly Leu Leu Ser Leu Leu Ile Phe Ser Met Val Leu Pro Ala His
20 25 30
Ala Gly Ile Val Ile Tyr Gly Thr Arg Ile Ile Tyr Pro Ala Glu Asn
35 40 45
Lys Glu Val Met Val Gln Leu Met Asn Gln Gly Asn Arg Ser Ser Leu
50 55 60
Leu Gln Ala Trp Ile Asp Asp Gly Asp Thr Ser Leu Pro Pro Glu Lys
65 70 75 80
Ile Gln Val Pro Phe Met Leu Thr Pro Pro Val Ala Lys Ile Gly Ala
85 90 95
Asn Ser Gly Gln Gln Val Lys Ile Lys Ile Met Pro Asn Lys Leu Pro
100 105 110
Thr Asn Lys Glu Ser Ile Phe Tyr Leu Asn Val Leu Asp Ile Pro Pro
115 120 125
Asn Ser Pro Glu Gln Glu Gly Lys Asn Ala Leu Lys Phe Ala Met Gln
130 135 140
Asn Arg Ile Lys Leu Phe Tyr Arg Pro Ala Gly Ile Ala Pro Val Asn
145 150 155 160
Lys Ala Thr Phe Lys Lys Leu Leu Val Asn Arg Ser Gly Asn Gly Leu
165 170 175
Val Ile Lys Asn Asp Ser Ala Asn Trp Val Thr Ile Ser Asp Val Lys
180 185 190
Ala Asn Asn Val Lys Val Asn Tyr Glu Thr Ile Met Ile Ala Pro Leu
195 200 205
Glu Ser Gln Ser Val Asn Val Lys Ser Asn Asn Ala Asn Asn Trp His
210 215 220
Leu Thr Ile Ile Asp Asp His Gly Asn Tyr Ile Ser Asp Lys Ile
225 230 235

<210> 311
<211> 180
<212> PRT
<213> Escherichia coli

<400> 311
Met Lys Arg Ser Ile Ile Ala Ala Ala Val Phe Ser Ser Phe Phe Met
1 5 10 15
Ser Ala Gly Val Phe Ala Ala Asp Val Asp Thr Gly Thr Leu Thr Ile
20 25 30
Lys Gly Asn Ile Ala Glu Ser Pro Cys Lys Phe Glu Ala Gly Gly Asp
35 40 45
Ser Val Ser Ile Asn Met Pro Thr Val Pro Thr Ser Val Phe Glu Gly
50 55 60
Lys Ala Lys Tyr Ser Thr Tyr Asp Asp Ala Val Gly Val Thr Ser Ser
65 70 75 80
Met Leu Lys Ile Ser Cys Pro Lys Glu Val Ala Gly Val Lys Leu Ser

85 90 95
Leu Ile Thr Asn Asp Lys Ile Thr Gly Asn Asp Lys Ala Ile Ala Ser
100 105 110
Ser Asn Asp Thr Val Gly Tyr Tyr Leu Tyr Leu Gly Asp Asn Ser Asp
115 120 125
Val Leu Asp Val Ser Ala Pro Phe Asn Ile Glu Ser Tyr Lys Thr Ala
130 135 140
Glu Gly Gln Tyr Ala Ile Pro Phe Lys Ala Lys Tyr Leu Lys Leu Thr
145 150 155 160
Asp Asn Ser Val Gln Ser Gly Asp Val Leu Ser Ser Leu Val Met Arg
165 170 175
Val Ala Gln Asp
180

<210> 312
<211> 95
<212> PRT
<213> Escherichia coli

<400> 312
Met Leu His Thr Leu His Arg Ser Pro Trp Leu Thr Asp Phe Ala Ala
1 5 10 15
Leu Leu Arg Leu Leu Ser Glu Gly Asp Glu Leu Leu Leu Leu Gln Asp
20 25 30
Gly Val Thr Ala Ala Val Asp Gly Asn Arg Tyr Leu Glu Ser Leu Arg
35 40 45
Asn Ala Pro Ile Lys Val Tyr Ala Leu Asn Glu Asp Leu Ile Ala Arg
50 55 60
Gly Leu Thr Gly Gln Ile Ser Asn Asp Ile Ile Leu Ile Asp Tyr Thr
65 70 75 80
Asp Phe Val Arg Leu Thr Val Lys His Pro Ser Gln Met Ala Trp
85 90 95

<210> 313
<211> 119
<212> PRT
<213> Escherichia coli

<400> 313
Met Lys Arg Ile Ala Phe Val Phe Ser Thr Ala Pro His Gly Thr Ala
1 5 10 15
Ala Gly Arg Glu Gly Leu Asp Ala Leu Leu Ala Thr Ser Ala Leu Thr
20 25 30
Asp Asp Leu Ala Val Phe Phe Ile Ala Asp Gly Val Phe Gln Leu Leu
35 40 45
Pro Gly Gln Lys Pro Asp Ala Val Leu Ala Arg Asp Tyr Ile Ala Thr
50 55 60
Phe Lys Leu Leu Gly Leu Tyr Asp Ile Glu Gln Cys Trp Val Cys Ala
65 70 75 80
Ala Ser Leu Arg Glu Arg Gly Leu Asp Pro Gln Thr Pro Phe Val Val
85 90 95
Glu Ala Thr Pro Leu Glu Ala Asp Ala Leu Arg Arg Glu Leu Ala Asn
100 105 110
Tyr Asp Val Ile Leu Arg Phe
115

<210> 314
<211> 128

<212> PRT
<213> Escherichia coli

<400> 314
Met Arg Phe Ala Ile Val Val Thr Gly Pro Ala Tyr Gly Thr Gln Gln
1 5 10 15
Ala Ser Ser Ala Phe Gln Phe Ala Gln Ala Leu Ile Ala Asp Gly His
20 25 30
Glu Leu Ser Ser Val Phe Phe Tyr Arg Glu Gly Val Tyr Asn Ala Asn
35 40 45
Gln Leu Thr Ser Pro Ala Ser Asp Glu Phe Asp Leu Val Arg Ala Trp
50 55 60
Gln Gln Leu Asn Ala Gln His Gly Val Ala Leu Asn Ile Cys Val Ala
65 70 75 80
Ala Ala Leu Arg Arg Gly Val Val Asp Glu Thr Glu Ala Gly Arg Leu
85 90 95
Gly Leu Ala Ser Ser Asn Leu Gln Gln Gly Phe Thr Leu Ser Gly Leu
100 105 110
Gly Ala Leu Ala Glu Ala Ser Leu Thr Cys Asp Arg Val Val Gln Phe
115 120 125

<210> 315
<211> 244
<212> PRT
<213> Escherichia coli

<400> 315
Met Phe Phe Phe Met Ser Arg Ser Leu Leu Thr Asn Glu Thr Ser Glu
1 5 10 15
Leu Asp Leu Leu Asp Gln Arg Pro Phe Asp Gln Thr Asp Phe Asp Ile
20 25 30
Leu Lys Ser Tyr Glu Ala Val Val Asp Gly Leu Ala Met Leu Ile Gly
35 40 45
Ser His Cys Glu Ile Val Leu His Ser Leu Gln Asp Leu Lys Cys Ser
50 55 60
Ala Ile Arg Ile Ala Asn Gly Glu His Thr Gly Arg Lys Ile Gly Ser
65 70 75 80
Pro Ile Thr Asp Leu Ala Leu Arg Met Leu His Asp Met Thr Gly Ala
85 90 95
Asp Ser Ser Val Ser Lys Cys Tyr Phe Thr Arg Ala Lys Ser Gly Val
100 105 110
Leu Met Lys Ser Leu Thr Ile Ala Ile Arg Asn Arg Glu Gln Arg Val
115 120 125
Ile Gly Leu Leu Cys Ile Asn Met Asn Leu Asp Val Pro Phe Ser Gln
130 135 140
Ile Met Ser Thr Phe Val Pro Pro Glu Thr Pro Asp Val Gly Ser Ser
145 150 155 160
Val Asn Phe Ala Ser Ser Val Glu Asp Leu Val Thr Gln Thr Leu Glu
165 170 175
Phe Thr Ile Glu Glu Val Asn Ala Asp Arg Asn Val Ser Asn Asn Ala
180 185 190
Lys Asn Arg Gln Ile Val Leu Asn Leu Tyr Glu Lys Gly Ile Phe Asp
195 200 205
Ile Lys Asp Ala Ile Asn Gln Val Ala Asp Arg Leu Asn Ile Ser Lys
210 215 220
His Thr Val Tyr Leu Tyr Ile Arg Gln Phe Lys Ser Gly Asp Phe Gln
225 230 235 240
Gly Gln Asp Lys

<210> 316
<211> 84
<212> PRT
<213> Escherichia coli

<400> 316
Met Thr Asp Lys Ile Arg Thr Leu Gln Gly Arg Val Val Ser Asp Lys
1   5   10   15
Met Glu Lys Ser Ile Val Val Ala Ile Glu Arg Phe Val Lys His Pro
            20   25   30
Ile Tyr Gly Lys Phe Ile Lys Arg Thr Thr Lys Leu His Val His Asp
        35   40   45
Glu Asn Asn Glu Cys Gly Ile Gly Asp Val Val Glu Ile Arg Glu Cys
    50   55   60
Arg Pro Leu Ser Lys Thr Lys Ser Trp Thr Leu Val Arg Val Val Glu
65   70   75   80
Lys Ala Val Leu

<210> 317
<211> 63
<212> PRT
<213> Escherichia coli

<400> 317
Met Lys Ala Lys Glu Leu Arg Glu Lys Ser Val Glu Glu Leu Asn Thr
1   5   10   15
Glu Leu Leu Asn Leu Leu Arg Glu Gln Phe Asn Leu Arg Met Gln Ala
            20   25   30
Ala Ser Gly Gln Leu Gln Gln Ser His Leu Leu Lys Gln Val Arg Arg
        35   40   45
Asp Val Ala Arg Val Lys Thr Leu Leu Asn Glu Lys Ala Gly Ala
    50   55   60

<210> 318
<211> 136
<212> PRT
<213> Escherichia coli

<400> 318
Met Leu Gln Pro Lys Arg Thr Lys Phe Arg Lys Met His Lys Gly Arg
1   5   10   15
Asn Arg Gly Leu Ala Gln Gly Thr Asp Val Ser Phe Gly Ser Phe Gly
            20   25   30
Leu Lys Ala Val Gly Arg Gly Arg Leu Thr Ala Arg Gln Ile Glu Ala
        35   40   45
Ala Arg Arg Ala Met Thr Arg Ala Val Lys Arg Gln Gly Lys Ile Trp
    50   55   60
Ile Arg Val Phe Pro Asp Lys Pro Ile Thr Glu Lys Pro Leu Ala Val
65   70   75   80
Arg Met Gly Lys Gly Lys Gly Asn Val Glu Tyr Trp Val Ala Leu Ile
                85   90   95
Gln Pro Gly Lys Val Leu Tyr Glu Met Asp Gly Val Pro Glu Glu Leu
            100   105   110
Ala Arg Glu Ala Phe Lys Leu Ala Ala Ala Lys Leu Pro Ile Lys Thr
        115   120   125
Thr Phe Val Thr Lys Thr Val Met
    130   135

<210> 319
<211> 233
<212> PRT
<213> Escherichia coli

<400> 319
Met Gly Gln Lys Val His Pro Asn Gly Ile Arg Leu Gly Ile Val Lys
1   5   10   15
Pro Trp Asn Ser Thr Trp Phe Ala Asn Thr Lys Glu Phe Ala Asp Asn
            20   25   30
Leu Asp Ser Asp Phe Lys Val Arg Gln Tyr Leu Thr Lys Glu Leu Ala
        35   40   45
Lys Ala Ser Val Ser Arg Ile Val Ile Glu Arg Pro Ala Lys Ser Ile
    50   55   60
Arg Val Thr Ile His Thr Ala Arg Pro Gly Ile Val Ile Gly Lys Lys
65   70   75   80
Gly Glu Asp Val Glu Lys Leu Arg Lys Val Val Ala Asp Ile Ala Gly
                85   90   95
Val Pro Ala Gln Ile Asn Ile Ala Glu Val Arg Lys Pro Glu Leu Asp
            100   105   110
Ala Lys Leu Val Ala Asp Ser Ile Thr Ser Gln Leu Glu Arg Arg Val
        115   120   125
Met Phe Arg Arg Ala Met Lys Arg Ala Val Gln Asn Ala Met Arg Leu
    130   135   140
Gly Ala Lys Gly Ile Lys Val Glu Val Ser Gly Arg Leu Gly Gly Ala
145   150   155   160
Glu Ile Ala Arg Thr Glu Trp Tyr Arg Glu Gly Arg Val Pro Leu His
                165   170   175
Thr Leu Arg Ala Asp Ile Asp Tyr Asn Thr Ser Glu Ala His Thr Thr
            180   185   190
Tyr Gly Val Ile Gly Val Lys Val Trp Ile Phe Lys Gly Glu Ile Leu
        195   200   205
Gly Gly Met Ala Ala Val Glu Gln Pro Glu Lys Pro Ala Ala Gln Pro
    210   215   220
Lys Lys Gln Gln Arg Lys Gly Arg Lys
225   230

<210> 320
<211> 110
<212> PRT
<213> Escherichia coli

<400> 320
Met Glu Thr Ile Ala Lys His Arg His Ala Arg Ser Ser Ala Gln Lys
1   5   10   15
Val Arg Leu Val Ala Asp Leu Ile Arg Gly Lys Lys Val Ser Gln Ala
            20   25   30
Leu Asp Ile Leu Thr Tyr Thr Asn Lys Lys Ala Ala Val Leu Val Lys
        35   40   45
Lys Val Leu Glu Ser Ala Ile Ala Asn Ala Glu His Asn Asp Gly Ala
    50   55   60
Asp Ile Asp Asp Leu Lys Val Thr Lys Ile Phe Val Asp Glu Gly Pro
65   70   75   80
Ser Met Lys Arg Ile Met Pro Arg Ala Lys Gly Arg Ala Asp Arg Ile
                85   90   95
Leu Lys Arg Thr Ser His Ile Thr Val Val Val Ser Asp Arg
            100   105   110

<210> 321
<211> 92
<212> PRT
<213> Escherichia coli

<400> 321
Met Pro Arg Ser Leu Lys Lys Gly Pro Phe Ile Asp Leu His Leu Leu
1 5 10 15
Lys Lys Val Glu Lys Ala Val Glu Ser Gly Asp Lys Lys Pro Leu Arg
20 25 30
Thr Trp Ser Arg Arg Ser Thr Ile Phe Pro Asn Met Ile Gly Leu Thr
35 40 45
Ile Ala Val His Asn Gly Arg Gln His Val Pro Val Phe Val Thr Asp
50 55 60
Glu Met Val Gly His Lys Leu Gly Glu Phe Ala Pro Thr Arg Thr Tyr
65 70 75 80
Arg Gly His Ala Ala Asp Lys Lys Ala Lys Lys Lys
85 90

<210> 322
<211> 273
<212> PRT
<213> Escherichia coli

<400> 322
Met Ala Val Val Lys Cys Lys Pro Thr Ser Pro Gly Arg Arg His Val
1 5 10 15
Val Lys Val Val Asn Pro Glu Leu His Lys Gly Lys Pro Phe Ala Pro
20 25 30
Leu Leu Glu Lys Asn Ser Lys Ser Gly Gly Arg Asn Asn Asn Gly Arg
35 40 45
Ile Thr Thr Arg His Ile Gly Gly Gly His Lys Gln Ala Tyr Arg Ile
50 55 60
Val Asp Phe Lys Arg Asn Lys Asp Gly Ile Pro Ala Val Val Glu Arg
65 70 75 80
Leu Glu Tyr Asp Pro Asn Arg Ser Ala Asn Ile Ala Leu Val Leu Tyr
85 90 95
Lys Asp Gly Glu Arg Arg Tyr Ile Leu Ala Pro Lys Gly Leu Lys Ala
100 105 110
Gly Asp Gln Ile Gln Ser Gly Val Asp Ala Ala Ile Lys Pro Gly Asn
115 120 125
Thr Leu Pro Met Arg Asn Ile Pro Val Gly Ser Thr Val His Asn Val
130 135 140
Glu Met Lys Pro Gly Lys Gly Gly Gln Leu Ala Arg Ser Ala Gly Thr
145 150 155 160
Tyr Val Gln Ile Val Ala Arg Asp Gly Ala Tyr Val Thr Leu Arg Leu
165 170 175
Arg Ser Gly Glu Met Arg Lys Val Glu Ala Asp Cys Arg Ala Thr Leu
180 185 190
Gly Glu Val Gly Asn Ala Glu His Met Leu Arg Val Leu Gly Lys Ala
195 200 205
Gly Ala Ala Arg Trp Arg Gly Val Arg Pro Thr Val Arg Gly Thr Ala
210 215 220
Met Asn Pro Val Asp His Pro His Gly Gly Gly Glu Gly Arg Asn Phe
225 230 235 240
Gly Lys His Pro Val Thr Pro Trp Gly Val Gln Thr Lys Gly Lys Lys
245 250 255
Thr Arg Ser Asn Lys Arg Thr Asp Lys Phe Ile Val Arg Arg Arg Ser
260 265 270

Lys

<210> 323
<211> 100
<212> PRT
<213> Escherichia coli

<400> 323
Met Ile Arg Glu Glu Arg Leu Leu Lys Val Leu Arg Ala Pro His Val
1 5 10 15
Ser Glu Lys Ala Ser Thr Ala Met Glu Lys Ser Asn Thr Ile Val Leu
20 25 30
Lys Val Ala Lys Asp Ala Thr Lys Ala Glu Ile Lys Ala Ala Val Gln
35 40 45
Lys Leu Phe Glu Val Glu Val Glu Val Val Asn Thr Leu Val Val Lys
50 55 60
Gly Lys Val Lys Arg His Gly Gln Arg Ile Gly Arg Arg Ser Asp Trp
65 70 75 80
Lys Lys Ala Tyr Val Thr Leu Lys Glu Gly Gln Asn Leu Asp Phe Val
85 90 95
Gly Gly Ala Glu
100

<210> 324
<211> 201
<212> PRT
<213> Escherichia coli

<400> 324
Met Glu Leu Val Leu Lys Asp Ala Gln Ser Ala Leu Thr Val Ser Glu
1 5 10 15
Thr Thr Phe Gly Arg Asp Phe Asn Glu Ala Leu Val His Gln Val Val
20 25 30
Val Ala Tyr Ala Ala Gly Ala Arg Gln Gly Thr Arg Ala Gln Lys Thr
35 40 45
Arg Ala Glu Val Thr Gly Ser Gly Lys Lys Pro Trp Arg Gln Lys Gly
50 55 60
Thr Gly Arg Ala Arg Ser Gly Ser Ile Lys Ser Pro Ile Trp Arg Ser
65 70 75 80
Gly Gly Val Thr Phe Ala Ala Arg Pro Gln Asp His Ser Gln Lys Val
85 90 95
Asn Lys Lys Met Tyr Arg Gly Ala Leu Lys Ser Ile Leu Ser Glu Leu
100 105 110
Val Arg Gln Asp Arg Leu Ile Val Val Glu Lys Phe Ser Val Glu Ala
115 120 125
Pro Lys Thr Lys Leu Leu Ala Gln Lys Leu Lys Asp Met Ala Leu Glu
130 135 140
Asp Val Leu Ile Ile Thr Gly Glu Leu Asp Glu Asn Leu Phe Leu Ala
145 150 155 160
Ala Arg Asn Leu His Lys Val Asp Val Arg Asp Ala Thr Gly Ile Asp
165 170 175
Pro Val Ser Leu Ile Ala Phe Asp Lys Val Val Met Thr Ala Asp Ala
180 185 190
Val Lys Gln Val Glu Glu Met Leu Ala
195 200

<210> 325

<211> 209
<212> PRT
<213> Escherichia coli

<400> 325
Met Ile Gly Leu Val Gly Lys Lys Val Gly Met Thr Arg Ile Phe Thr
1 5 10 15
Glu Asp Gly Val Ser Ile Pro Val Thr Val Ile Glu Val Glu Ala Asn
20 25 30
Arg Val Thr Gln Val Lys Asp Leu Ala Asn Asp Gly Tyr Arg Ala Ile
35 40 45
Gln Val Thr Thr Gly Ala Lys Lys Ala Asn Arg Val Thr Lys Pro Glu
50 55 60
Ala Gly His Phe Ala Lys Ala Gly Val Glu Ala Gly Arg Gly Leu Trp
65 70 75 80
Glu Phe Arg Leu Ala Glu Gly Glu Glu Phe Thr Val Gly Gln Ser Ile
85 90 95
Ser Val Glu Leu Phe Ala Asp Val Lys Lys Val Asp Val Thr Gly Thr
100 105 110
Ser Lys Gly Lys Gly Phe Ala Gly Thr Val Lys Arg Trp Asn Phe Arg
115 120 125
Thr Gln Asp Ala Thr His Gly Asn Ser Leu Ser His Arg Val Pro Gly
130 135 140
Ser Ile Gly Gln Asn Gln Thr Pro Gly Lys Val Phe Lys Gly Lys Lys
145 150 155 160
Met Ala Gly Gln Met Gly Asn Glu Arg Val Thr Val Gln Ser Leu Asp
165 170 175
Val Val Arg Val Asp Ala Glu Arg Asn Leu Leu Leu Val Lys Gly Ala
180 185 190
Val Pro Gly Ala Thr Gly Ser Asp Leu Ile Val Lys Pro Ala Val Lys
195 200 205
Ala

<210> 326
<211> 103
<212> PRT
<213> Escherichia coli

<400> 326
Met Gln Asn Gln Arg Ile Arg Ile Arg Leu Lys Ala Phe Asp His Arg
1 5 10 15
Leu Ile Asp Gln Ala Thr Ala Glu Ile Val Glu Thr Ala Lys Arg Thr
20 25 30
Gly Ala Gln Val Arg Gly Pro Ile Pro Leu Pro Thr Arg Lys Glu Arg
35 40 45
Phe Thr Val Leu Ile Ser Pro His Val Asn Lys Asp Ala Arg Asp Gln
50 55 60
Tyr Glu Ile Arg Thr His Leu Arg Leu Val Asp Ile Val Glu Pro Thr
65 70 75 80
Glu Lys Thr Val Asp Ala Leu Met Arg Leu Asp Leu Ala Ala Gly Val
85 90 95
Asp Val Gln Ile Ser Leu Gly
100

<210> 327
<211> 104
<212> PRT
<213> Escherichia coli

<400> 327
Met Ile Arg Lys Ala Phe Val Met Gln Val Asn Pro Asp Ala His Glu
1 5 10 15
Glu Tyr Gln Arg Arg His Asn Pro Ile Trp Pro Glu Leu Glu Ala Val
20 25 30
Leu Lys Ser His Gly Ala His Asn Tyr Ala Ile Tyr Leu Asp Lys Ala
35 40 45
Arg Asn Leu Leu Phe Ala Met Val Glu Ile Glu Ser Glu Glu Arg Trp
50 55 60
Asn Ala Val Ala Ser Thr Asp Val Cys Gln Arg Trp Trp Lys Tyr Met
65 70 75 80
Thr Asp Val Met Pro Ala Asn Pro Asp Asn Ser Pro Val Ser Ser Glu
85 90 95
Leu Gln Glu Val Phe Tyr Leu Pro
100

<210> 328
<211> 287
<212> PRT
<213> Escherichia coli

<400> 328
Met Ile Arg Ser Met Thr Ala Tyr Ala Arg Arg Glu Ile Lys Gly Glu
1 5 10 15
Trp Gly Ser Ala Thr Trp Glu Met Arg Ser Val Asn Gln Arg Tyr Leu
20 25 30
Glu Thr Tyr Phe Arg Leu Pro Glu Gln Phe Arg Ser Leu Glu Pro Val
35 40 45
Val Arg Glu Arg Ile Arg Ser Arg Leu Thr Arg Gly Lys Val Glu Cys
50 55 60
Thr Leu Arg Tyr Glu Pro Asp Val Ser Ala Gln Gly Glu Leu Ile Leu
65 70 75 80
Asn Glu Lys Leu Ala Lys Gln Leu Val Thr Ala Ala Asn Trp Val Lys
85 90 95
Met Gln Ser Asp Glu Gly Glu Ile Asn Pro Val Asp Ile Leu Arg Trp
100 105 110
Pro Gly Val Met Ala Ala Gln Glu Gln Asp Leu Asp Ala Ile Ala Ala
115 120 125
Glu Ile Leu Ala Ala Leu Asp Gly Thr Leu Asp Asp Phe Ile Val Ala
130 135 140
Arg Glu Thr Glu Gly Gln Ala Leu Lys Ala Leu Ile Glu Gln Arg Leu
145 150 155 160
Glu Gly Val Thr Ala Glu Val Val Lys Val Arg Ser His Met Pro Glu
165 170 175
Ile Leu Gln Trp Gln Arg Glu Arg Leu Val Ala Lys Leu Glu Asp Ala
180 185 190
Gln Val Gln Leu Glu Asn Asn Arg Leu Glu Gln Glu Leu Val Leu Leu
195 200 205
Ala Gln Arg Ile Asp Val Ala Glu Glu Leu Asp Arg Leu Glu Ala His
210 215 220
Val Lys Glu Thr Tyr Asn Ile Leu Lys Lys Lys Glu Ala Val Gly Arg
225 230 235 240
Arg Leu Asp Phe Met Met Gln Glu Phe Asn Arg Glu Ser Asn Thr Leu
245 250 255
Ala Ser Lys Ser Ile Asn Ala Glu Val Thr Asn Ser Ala Ile Glu Leu
260 265 270
Lys Val Leu Ile Glu Gln Met Arg Glu Gln Ile Gln Asn Ile Glu
275 280 285

<210> 329
<211> 163
<212> PRT
<213> Escherichia coli

<400> 329
Met Ser Thr Glu Thr Ile Glu Ile Phe Asn Asn Ser Asp Glu Trp Ala
1 5 10 15
Asn Gln Leu Lys His Ala Leu Ser Lys Gly Glu Asn Leu Ala Leu Leu
20 25 30
His Gly Leu Thr Pro Asp Ile Leu Asp Arg Ile Tyr Ala Tyr Ala Phe
35 40 45
Asp Tyr His Glu Lys Gly Asn Ile Thr Asp Ala Glu Ile Tyr Tyr Lys
50 55 60
Phe Leu Cys Ile Tyr Ala Phe Glu Asn His Glu Tyr Leu Lys Asp Phe
65 70 75 80
Ala Ser Val Cys Gln Pro Lys Lys Lys Tyr Gln Gln Ala Tyr Asp Leu
85 90 95
Tyr Lys Leu Ser Tyr Asn Tyr Phe Pro Tyr Asp Asp Tyr Ser Val Ile
100 105 110
Tyr Arg Met Gly Gln Cys Gln Ile Gly Ala Lys Asn Ile Asp Asn Ala
115 120 125
Met Gln Cys Phe Tyr His Ile Ile Asn Asn Cys Glu Asp Asp Ser Val
130 135 140
Lys Ser Lys Ala Gln Ala Tyr Ile Glu Leu Leu Asn Asp Asn Ser Glu
145 150 155 160
Asp Asn Gly

<210> 330
<211> 648
<212> PRT
<213> Escherichia coli

<400> 330
Met Asn Ile Leu Gly Phe Phe Gln Arg Leu Gly Arg Ala Leu Gln Leu
1 5 10 15
Pro Ile Ala Val Leu Pro Val Ala Ala Leu Leu Leu Arg Phe Gly Gln
20 25 30
Pro Asp Leu Leu Asn Val Ala Phe Ile Ala Gln Ala Gly Gly Ala Ile
35 40 45
Phe Asp Asn Leu Ala Leu Ile Phe Ala Ile Gly Val Ala Ser Ser Trp
50 55 60
Ser Lys Asp Ser Ala Gly Ala Ala Ala Leu Ala Gly Ala Val Gly Tyr
65 70 75 80
Phe Val Leu Thr Lys Ala Met Val Thr Ile Asn Pro Glu Ile Asn Met
85 90 95
Gly Val Leu Ala Gly Ile Ile Thr Gly Leu Val Gly Gly Ala Ala Tyr
100 105 110
Asn Arg Trp Ser Asp Ile Lys Leu Pro Asp Phe Leu Ser Phe Phe Gly
115 120 125
Gly Lys Arg Phe Val Pro Ile Ala Thr Gly Phe Phe Cys Leu Val Leu
130 135 140
Ala Ala Ile Phe Gly Tyr Val Trp Pro Pro Val Gln His Ala Ile His
145 150 155 160
Ala Gly Gly Glu Trp Ile Val Ser Ala Gly Ala Leu Gly Ser Gly Ile
165 170 175
Phe Gly Phe Ile Asn Arg Leu Leu Ile Pro Thr Gly Leu His Gln Val
180 185 190

Leu Asn Thr Ile Ala Trp Phe Gln Ile Gly Glu Phe Thr Asn Ala Ala
195 200 205
Gly Thr Val Phe His Gly Asp Ile Asn Arg Phe Tyr Ala Gly Asp Gly
210 215 220
Thr Ala Gly Met Phe Met Ser Gly Phe Phe Pro Ile Met Met Phe Gly
225 230 235 240
Leu Pro Gly Ala Ala Leu Ala Met Tyr Phe Ala Ala Pro Lys Glu Arg
245 250 255
Arg Pro Met Val Gly Gly Met Leu Leu Ser Val Ala Val Thr Ala Phe
260 265 270
Leu Thr Gly Val Thr Glu Pro Leu Glu Phe Leu Phe Met Phe Leu Ala
275 280 285
Pro Leu Leu Tyr Leu Leu His Ala Leu Leu Thr Gly Ile Ser Leu Phe
290 295 300
Val Ala Thr Leu Leu Gly Ile His Ala Gly Phe Ser Phe Ser Ala Gly
305 310 315 320
Ala Ile Asp Tyr Ala Leu Met Tyr Asn Leu Pro Ala Ala Ser Gln Asn
325 330 335
Val Trp Met Leu Leu Val Met Gly Val Ile Phe Phe Ala Ile Tyr Phe
340 345 350
Val Val Phe Ser Leu Val Ile Arg Met Phe Asn Leu Lys Thr Pro Gly
355 360 365
Arg Glu Asp Lys Glu Asp Glu Ile Val Thr Glu Glu Ala Asn Ser Asn
370 375 380
Thr Glu Glu Gly Leu Thr Gln Leu Ala Thr Asn Tyr Ile Ala Ala Val
385 390 395 400
Gly Gly Thr Asp Asn Leu Lys Ala Ile Asp Ala Cys Ile Thr Arg Leu
405 410 415
Arg Leu Thr Val Ala Asp Ser Ala Arg Val Asn Asp Thr Met Cys Lys
420 425 430
Arg Leu Gly Ala Ser Gly Val Val Lys Leu Asn Lys Gln Thr Ile Gln
435 440 445
Val Ile Val Gly Ala Lys Ala Glu Ser Ile Gly Asp Ala Met Lys Lys
450 455 460
Val Val Ala Arg Gly Pro Val Ala Ala Ala Ser Ala Glu Ala Thr Pro
465 470 475 480
Ala Thr Ala Ala Pro Val Ala Lys Pro Gln Ala Val Pro Asn Ala Val
485 490 495
Ser Ile Ala Glu Leu Val Ser Pro Ile Thr Gly Asp Val Val Ala Leu
500 505 510
Asp Gln Val Pro Asp Glu Ala Phe Ala Ser Lys Ala Val Gly Asp Gly
515 520 525
Val Ala Val Lys Pro Thr Asp Lys Ile Val Val Ser Pro Ala Ala Gly
530 535 540
Thr Ile Val Lys Ile Phe Asn Thr Asn His Ala Phe Cys Leu Glu Thr
545 550 555 560
Glu Lys Gly Ala Glu Ile Val Val His Met Gly Ile Asp Thr Val Ala
565 570 575
Leu Glu Gly Lys Gly Phe Lys Arg Leu Val Glu Glu Gly Ala Gln Val
580 585 590
Ser Ala Gly Gln Pro Ile Leu Glu Met Asp Leu Asp Tyr Leu Asn Ala
595 600 605
Asn Ala Arg Ser Met Ile Ser Pro Val Val Cys Ser Asn Ile Asp Asp
610 615 620
Phe Ser Gly Leu Ile Ile Lys Ala Gln Gly His Ile Val Ala Gly Gln
625 630 635 640
Thr Pro Leu Tyr Glu Ile Lys Lys
645

<210> 331

<211> 412
<212> PRT
<213> Escherichia coli

<400> 331
Met Lys Thr Ile Phe Arg Tyr Ile Leu Phe Leu Ala Leu Tyr Ser Cys
1 5 10 15
Cys Asn Thr Val Ser Ala Tyr Thr Ser Phe Ile Val Gly Asn Asn Ala
20 25 30
Gly Val Asp Asn Tyr Arg Gly Pro Ser Thr Ala Ala Gln Met Thr Phe
35 40 45
Asn Tyr Thr Ser Thr Ala Ser Asn Leu Val Phe Tyr Lys Pro Thr Gln
50 55 60
Leu Gly Pro Thr Gly Val Lys Met Tyr Trp Ser Tyr Leu Asp Thr Gly
65 70 75 80
Thr Gly Gly Gly Ile Leu Tyr Cys Asn Thr Ser Gly Arg Ala Asn Pro
85 90 95
Gly Pro Ile Thr Ile Glu Asn Ala Met Val Tyr Ser Gly Lys Asp Tyr
100 105 110
Gly Gly His Lys Leu Phe Asn Thr Ser Val Pro Gly Leu Tyr Tyr Thr
115 120 125
Met Leu Ile Ser Arg Val Trp Ser Ala Tyr Asp Thr Ile Thr Asp Ile
130 135 140
Gln Ser Pro Gly Ile Tyr Ile Gly Asp Pro Ser Asn Gln Glu Phe Phe
145 150 155 160
Phe Ser Val Thr Asp Ser Asp Leu Gln Thr Lys Gly Cys Asn Lys Ala
165 170 175
Asp Asp Tyr Asp Lys Phe Trp Ala Ile Gly Gly Ile Val His Asn Ile
180 185 190
Thr Val Glu Phe Tyr Thr Asp Thr Asn Phe Asp Pro Thr Leu Asn Gln
195 200 205
Gln Val Gln Leu Ser Ser Ser Ser Asn Tyr Leu Tyr Ser Phe Lys Ala
210 215 220
Tyr Ser Pro Gly Thr Lys Val Val Asp His Ser Asn His Ile Tyr Val
225 230 235 240
Asn Phe Thr Leu Asn Asn Val Lys Leu Thr Leu Pro Thr Cys Phe Thr
245 250 255
Ser Ile Leu Thr Gly Pro Ser Val Asn Gly Ser Thr Val Arg Met Gly
260 265 270
Glu Tyr Ser Ser Gly Thr Ile Lys Asn Gly Ala Ser Pro Val Pro Phe
275 280 285
Asp Ile Ser Leu Gln Asn Cys Ile Arg Val Arg Asn Ile Glu Thr Lys
290 295 300
Leu Val Thr Gly Lys Val Gly Thr Gln Asn Thr Gln Leu Leu Gly Asn
305 310 315 320
Thr Leu Thr Gly Ser Thr Ala Ala Lys Gly Val Gly Val Leu Ile Glu
325 330 335
Gly Leu Ala Thr Ser Lys Asn Pro Leu Met Thr Leu Lys Pro Asn Asp
340 345 350
Thr Asn Ser Val Tyr Ile Asp Tyr Glu Thr Glu Asp Asp Thr Ser Asp
355 360 365
Gly Val Tyr Pro Asn Gln Gly Asn Gly Thr Ser Gln Pro Leu His Phe
370 375 380
Gln Ala Thr Leu Lys Gln Asp Gly Asn Ile Ala Ile Glu Pro Gly Glu
385 390 395 400
Phe Lys Ala Thr Ser Thr Phe Gln Val Thr Tyr Pro
405 410

<210> 332
<211> 198

<212> PRT
<213> Escherichia coli

<400> 332
Met His Pro Thr Gln Arg Lys Leu Met Lys Arg Ile Ile Leu Phe Leu
1 5 10 15
Ser Leu Leu Phe Cys Ile Ala Cys Pro Ala Ile Ala Gly Gln Asp Ile
20 25 30
Asp Leu Val Ala Asn Val Lys Asn Ser Thr Cys Lys Ser Gly Ile Ser
35 40 45
Asn Gln Gly Asn Ile Asp Leu Gly Val Val Gly Val Gly Tyr Phe Ser
50 55 60
Gly Asn Val Thr Pro Glu Ser Tyr Gln Pro Gly Gly Lys Glu Phe Thr
65 70 75 80
Ile Thr Val Ser Asp Cys Ala Leu Gln Gly Thr Gly Asp Val Leu Asn
85 90 95
Gln Leu His Ile Asp Phe Arg Ala Leu Ser Gly Val Met Ala Ala Gly
100 105 110
Ser Arg Gln Ile Phe Ala Asn Glu Ile Ser Ser Gly Ala Ser Asn Val
115 120 125
Gly Val Val Ile Phe Ser Thr Gln Asp Ser Ala Asn Thr Phe Asn Val
130 135 140
Leu Asn Ala Ser Gly Gly Ser Arg Ser Val Tyr Pro Val Met Ser Asp
145 150 155 160
Asp Met Asn Gly Ser Ser Trp Lys Phe Ser Thr Arg Met Gln Lys Ile
165 170 175
Asp Pro Ala Leu Ser Val Thr Ser Gly Gln Leu Met Ser His Val Leu
180 185 190
Val Asp Ile Tyr Tyr Glu
195

<210> 333
<211> 201
<212> PRT
<213> Escherichia coli

<400> 333
Met Met Thr Phe Lys Asn Leu Arg Tyr Gly Leu Ser Ser Ser Val Val
1 5 10 15
Leu Ala Ala Ser Leu Phe Ser Val Leu Ser Tyr Ala Ala Thr Asp Ser
20 25 30
Ile Gly Leu Thr Val Ile Thr Thr Val Glu Met Gly Thr Cys Thr Ala
35 40 45
Thr Leu Val Asn Asp Ser Asp Gln Asp Ile Ser Val Val Asp Phe Gly
50 55 60
Asp Val Tyr Ile Ser Glu Ile Asn Ala Lys Thr Lys Val Lys Thr Phe
65 70 75 80
Lys Leu Lys Phe Lys Asp Cys Ala Gly Ile Pro Asn Lys Lys Ala Gln
85 90 95
Ile Lys Leu Thr Lys Arg Ala Thr Cys Glu Gly Thr Ala Asn Asp Gly
100 105 110
Ala Gly Phe Ala Asn Gly Ser Thr Ala Ala Asp Lys Ala Ser Ala Val
115 120 125
Ala Val Glu Val Trp Ser Thr Val Thr Pro Ala Thr Gly Ser Ala Thr
130 135 140
Gln Phe Ser Cys Val Thr Pro Ala Ser Gln Glu Val Thr Ile Ser Thr
145 150 155 160
Ala Ala Asn Ala Val Val Tyr Tyr Pro Met Ser Ala Arg Leu Val Val
165 170 175
Glu Lys Asn Lys Thr Val Asn Asn Val Thr Ala Gly Lys Phe Ser Ala

```
        180                 185                 190
Pro Ala Thr Phe Thr Val Thr Tyr Asn
        195                 200


<210> 334
<211> 203
<212> PRT
<213> Escherichia coli

<400> 334
Met Glu Phe Gly Val Arg Phe Ser Asn Tyr Lys Gly Arg Gln Met Ile
 1               5                   10                  15
Lys Thr Thr Pro His Lys Ile Val Ile Leu Met Gly Ile Leu Leu Ser
            20                  25                  30
Pro Ser Val Phe Ala Thr Asp Ile Asn Val Glu Phe Thr Ala Thr Val
        35                  40                  45
Lys Ala Thr Thr Cys Asn Ile Thr Leu Thr Gly Asn Asn Val Thr Asn
    50                  55                  60
Asp Gly Asn Asn Asn Tyr Thr Leu Arg Ile Pro Lys Met Gly Leu Asp
65                  70                  75                  80
Lys Ile Ala Asn Lys Thr Thr Glu Ser Gln Ala Asp Phe Lys Leu Val
                85                  90                  95
Ala Ser Gly Cys Ser Ser Gly Ile Ser Trp Ile Asp Thr Thr Leu Thr
            100                 105                 110
Gly Asn Ala Ser Ser Ser Ser Pro Lys Leu Ile Ile Pro Gln Ser Gly
        115                 120                 125
Asp Ser Ser Ser Thr Thr Ser Asn Ile Gly Met Gly Phe Lys Lys Arg
    130                 135                 140
Thr Thr Asp Asp Ala Thr Phe Leu Lys Pro Asn Ser Ala Glu Lys Ile
145                 150                 155                 160
Arg Trp Ser Thr Asp Glu Met Gln Pro Asp Lys Gly Leu Glu Met Thr
                165                 170                 175
Val Ala Leu Arg Glu Thr Asp Ala Gly Gln Gly Val Pro Gly Asn Phe
            180                 185                 190
Arg Ala Leu Ala Thr Phe Asn Phe Ile Tyr Gln
        195                 200


<210> 335
<211> 139
<212> PRT
<213> Escherichia coli

<400> 335
Met Ala Met Thr Tyr His Leu Asp Val Val Ser Ala Glu Gln Gln Met
 1               5                   10                  15
Phe Ser Gly Leu Val Glu Lys Ile Gln Val Thr Gly Ser Glu Gly Glu
            20                  25                  30
Leu Gly Ile Tyr Pro Gly His Ala Pro Leu Leu Thr Ala Ile Lys Pro
        35                  40                  45
Gly Met Ile Arg Ile Val Lys Gln His Gly His Glu Glu Phe Ile Tyr
    50                  55                  60
Leu Ser Gly Gly Ile Leu Glu Val Gln Pro Gly Asn Val Thr Val Leu
65                  70                  75                  80
Ala Asp Thr Ala Ile Arg Gly Gln Asp Leu Asp Glu Ala Arg Ala Met
                85                  90                  95
Glu Ala Lys Arg Lys Ala Glu Glu His Ile Ser Ser Ser His Gly Asp
            100                 105                 110
Val Asp Tyr Ala Gln Ala Ser Ala Glu Leu Ala Lys Ala Ile Ala Gln
        115                 120                 125
```

Leu Arg Val Ile Glu Leu Thr Lys Lys Ala Met
130 135

<210> 336
<211> 460
<212> PRT
<213> Escherichia coli

<400> 336
Met Ala Thr Gly Lys Ile Val Gln Val Ile Gly Ala Val Val Asp Val
1 5 10 15
Glu Phe Pro Gln Asp Ala Val Pro Arg Val Tyr Asp Ala Leu Glu Val
20 25 30
Gln Asn Gly Asn Glu Arg Leu Val Leu Glu Val Gln Gln Leu Gly
35 40 45
Gly Gly Ile Val Arg Thr Ile Ala Met Gly Ser Ser Asp Gly Leu Arg
50 55 60
Arg Gly Leu Asp Val Lys Asp Leu Glu His Pro Ile Glu Val Pro Val
65 70 75 80
Gly Lys Ala Thr Leu Gly Arg Ile Met Asn Val Leu Gly Glu Pro Val
85 90 95
Asp Met Lys Gly Glu Ile Gly Glu Glu Glu Arg Trp Ala Ile His Arg
100 105 110
Ala Ala Pro Ser Tyr Glu Glu Leu Ser Asn Ser Gln Glu Leu Leu Glu
115 120 125
Thr Gly Ile Lys Val Ile Asp Leu Met Cys Pro Phe Ala Lys Gly Gly
130 135 140
Lys Val Gly Leu Phe Gly Gly Ala Gly Val Gly Lys Thr Val Asn Met
145 150 155 160
Met Glu Leu Ile Arg Asn Ile Ala Ile Glu His Ser Gly Tyr Ser Val
165 170 175
Phe Ala Gly Val Gly Glu Arg Thr Arg Glu Gly Asn Asp Phe Tyr His
180 185 190
Glu Met Thr Asp Ser Asn Val Ile Asp Lys Val Ser Leu Val Tyr Gly
195 200 205
Gln Met Asn Glu Pro Pro Gly Asn Arg Leu Arg Val Ala Leu Thr Gly
210 215 220
Leu Thr Met Ala Glu Lys Phe Arg Asp Glu Gly Arg Asp Val Leu Leu
225 230 235 240
Phe Val Asp Asn Ile Tyr Arg Tyr Thr Leu Ala Gly Thr Glu Val Ser
245 250 255
Ala Leu Leu Gly Arg Met Pro Ser Ala Val Gly Tyr Gln Pro Thr Leu
260 265 270
Ala Glu Glu Met Gly Val Leu Gln Glu Arg Ile Thr Ser Thr Lys Thr
275 280 285
Gly Ser Ile Thr Ser Val Gln Ala Val Tyr Val Pro Ala Asp Asp Leu
290 295 300
Thr Asp Pro Ser Pro Ala Thr Thr Phe Ala His Leu Asp Ala Thr Val
305 310 315 320
Val Leu Ser Arg Gln Ile Ala Ser Leu Gly Ile Tyr Pro Ala Val Asp
325 330 335
Pro Leu Asp Ser Thr Ser Arg Gln Leu Asp Pro Leu Val Val Gly Gln
340 345 350
Glu His Tyr Asp Thr Ala Arg Gly Val Gln Ser Ile Leu Gln Arg Tyr
355 360 365
Gln Glu Leu Lys Asp Ile Ile Ala Ile Leu Gly Met Asp Glu Leu Ser
370 375 380
Glu Glu Asp Lys Leu Val Val Ala Arg Ala Arg Lys Ile Gln Arg Phe
385 390 395 400
Leu Ser Gln Pro Phe Phe Val Ala Glu Val Phe Thr Gly Ser Pro Gly

405 410 415
Lys Tyr Val Ser Leu Lys Asp Thr Ile Arg Gly Phe Lys Gly Ile Met
420 425 430
Glu Gly Glu Tyr Asp His Leu Pro Glu Gln Ala Phe Tyr Met Val Gly
435 440 445
Ser Ile Glu Glu Ala Val Glu Lys Ala Lys Lys Leu
450 455 460

<210> 337
<211> 287
<212> PRT
<213> Escherichia coli

<400> 337
Met Ala Gly Ala Lys Glu Ile Arg Ser Lys Ile Ala Ser Val Gln Asn
1 5 10 15
Thr Gln Lys Ile Thr Lys Ala Met Glu Met Val Ala Ala Ser Lys Met
20 25 30
Arg Lys Ser Gln Asp Arg Met Ala Ala Ser Arg Pro Tyr Ala Glu Thr
35 40 45
Met Arg Lys Val Ile Gly His Leu Ala His Gly Asn Leu Glu Tyr Lys
50 55 60
His Pro Tyr Leu Glu Asp Arg Asp Val Lys Arg Val Gly Tyr Leu Val
65 70 75 80
Val Ser Thr Asp Arg Gly Leu Cys Gly Gly Leu Asn Ile Asn Leu Phe
85 90 95
Lys Lys Leu Leu Ala Glu Met Lys Thr Trp Thr Asp Lys Gly Val Gln
100 105 110
Cys Asp Leu Ala Met Ile Gly Ser Lys Gly Val Ser Phe Phe Asn Ser
115 120 125
Val Gly Gly Asn Val Val Ala Gln Val Thr Gly Met Gly Asp Asn Pro
130 135 140
Ser Leu Ser Glu Leu Ile Gly Pro Val Lys Val Met Leu Gln Ala Tyr
145 150 155 160
Asp Glu Gly Arg Leu Asp Lys Leu Tyr Ile Val Ser Asn Lys Phe Ile
165 170 175
Asn Thr Met Ser Gln Val Pro Thr Ile Ser Gln Leu Leu Pro Leu Pro
180 185 190
Ala Ser Asp Asp Asp Asp Leu Lys His Lys Ser Trp Asp Tyr Leu Tyr
195 200 205
Glu Pro Asp Pro Lys Ala Leu Leu Asp Thr Leu Leu Arg Arg Tyr Val
210 215 220
Glu Ser Gln Val Tyr Gln Gly Val Val Glu Asn Leu Ala Ser Glu Gln
225 230 235 240
Ala Ala Arg Met Val Ala Met Lys Ala Ala Thr Asp Asn Gly Gly Ser
245 250 255
Leu Ile Lys Glu Leu Gln Leu Val Tyr Asn Lys Ala Arg Gln Ala Ser
260 265 270
Ile Thr Gln Glu Leu Thr Glu Ile Val Ser Gly Ala Ala Ala Val
275 280 285

<210> 338
<211> 513
<212> PRT
<213> Escherichia coli

<400> 338
Met Gln Leu Asn Ser Thr Glu Ile Ser Glu Leu Ile Lys Gln Arg Ile
1 5 10 15

Ala Gln Phe Asn Val Val Ser Glu Ala His Asn Glu Gly Thr Ile Val
20 25 30
Ser Val Ser Asp Gly Val Ile Arg Ile His Gly Leu Ala Asp Cys Met
35 40 45
Gln Gly Glu Met Ile Ser Leu Pro Gly Asn Arg Tyr Ala Ile Ala Leu
50 55 60
Asn Leu Glu Arg Asp Ser Val Gly Ala Val Val Met Gly Pro Tyr Ala
65 70 75 80
Asp Leu Ala Glu Gly Met Lys Val Lys Cys Thr Gly Arg Ile Leu Glu
85 90 95
Val Pro Val Gly Arg Gly Leu Leu Gly Arg Val Val Asn Thr Leu Gly
100 105 110
Ala Pro Ile Asp Gly Lys Gly Pro Leu Asp His Asp Gly Phe Ser Ala
115 120 125
Val Glu Ala Ile Ala Pro Gly Val Ile Glu Arg Gln Ser Val Asp Gln
130 135 140
Pro Val Gln Thr Gly Tyr Lys Ala Val Asp Ser Met Ile Pro Ile Gly
145 150 155 160
Arg Gly Gln Arg Glu Leu Ile Ile Gly Asp Arg Gln Thr Gly Lys Thr
165 170 175
Ala Leu Ala Ile Asp Ala Ile Ile Asn Gln Arg Asp Ser Gly Ile Lys
180 185 190
Cys Ile Tyr Val Ala Ile Gly Gln Lys Ala Ser Thr Ile Ser Asn Val
195 200 205
Val Arg Lys Leu Glu Glu His Gly Ala Leu Ala Asn Thr Ile Val Val
210 215 220
Val Ala Thr Ala Ser Glu Ser Ala Ala Leu Gln Tyr Leu Ala Pro Tyr
225 230 235 240
Ala Gly Cys Ala Met Gly Glu Tyr Phe Arg Asp Arg Gly Glu Asp Ala
245 250 255
Leu Ile Ile Tyr Asp Asp Leu Ser Lys Gln Ala Val Ala Tyr Arg Gln
260 265 270
Ile Ser Leu Leu Leu Arg Arg Pro Pro Gly Arg Glu Ala Phe Pro Gly
275 280 285
Asp Val Phe Tyr Leu His Ser Arg Leu Leu Glu Arg Ala Ala Arg Val
290 295 300
Asn Ala Glu Tyr Val Glu Ala Phe Thr Lys Gly Glu Val Lys Gly Lys
305 310 315 320
Thr Gly Ser Leu Thr Ala Leu Pro Ile Ile Glu Thr Gln Ala Gly Asp
325 330 335
Val Ser Ala Phe Val Pro Thr Asn Val Ile Ser Ile Thr Asp Gly Gln
340 345 350
Ile Phe Leu Glu Thr Asn Leu Phe Asn Ala Gly Ile Arg Pro Ala Val
355 360 365
Asn Pro Gly Ile Ser Val Ser Arg Val Gly Gly Ala Ala Gln Thr Lys
370 375 380
Ile Met Lys Lys Leu Ser Gly Gly Ile Arg Thr Ala Leu Ala Gln Tyr
385 390 395 400
Arg Glu Leu Ala Ala Phe Ser Gln Phe Ala Ser Asp Leu Asp Asp Ala
405 410 415
Thr Arg Lys Gln Leu Asp His Gly Gln Lys Val Thr Glu Leu Leu Lys
420 425 430
Gln Lys Gln Tyr Ala Pro Met Ser Val Ala Gln Gln Ser Leu Val Leu
435 440 445
Phe Ala Ala Glu Arg Gly Tyr Leu Ala Asp Val Glu Leu Ser Lys Ile
450 455 460
Gly Ser Phe Glu Ala Ala Leu Leu Ala Tyr Val Asp Arg Asp His Ala
465 470 475 480
Pro Leu Met Gln Glu Ile Asn Gln Thr Gly Gly Tyr Asn Asp Glu Ile
485 490 495
Glu Gly Lys Leu Lys Gly Ile Leu Asp Ser Phe Lys Ala Thr Gln Ser

500 505 510
Trp

<210> 339
<211> 177
<212> PRT
<213> Escherichia coli

<400> 339
Met Ser Glu Phe Ile Thr Val Ala Arg Pro Tyr Ala Lys Ala Ala Phe
1 5 10 15
Asp Phe Ala Val Glu His Gln Ser Val Glu Arg Trp Gln Asp Met Leu
20 25 30
Ala Phe Ala Ala Glu Val Thr Lys Asn Glu Gln Met Ala Glu Leu Leu
35 40 45
Ser Gly Ala Leu Ala Pro Glu Thr Leu Ala Glu Ser Phe Ile Ala Val
50 55 60
Cys Gly Glu Gln Leu Asp Glu Asn Gly Gln Asn Leu Ile Arg Val Met
65 70 75 80
Ala Glu Asn Gly Arg Leu Asn Ala Leu Pro Asp Val Leu Glu Gln Phe
85 90 95
Ile His Leu Arg Ala Val Ser Glu Ala Thr Ala Glu Val Asp Val Ile
100 105 110
Ser Ala Ala Ala Leu Ser Glu Gln Gln Leu Ala Lys Ile Ser Ala Ala
115 120 125
Met Glu Lys Arg Leu Ser Arg Lys Val Lys Leu Asn Cys Lys Ile Asp
130 135 140
Lys Ser Val Met Ala Gly Val Ile Ile Arg Ala Gly Asp Met Val Ile
145 150 155 160
Asp Gly Ser Val Arg Gly Arg Leu Glu Arg Leu Ala Asp Val Leu Gln
165 170 175
Ser

<210> 340
<211> 156
<212> PRT
<213> Escherichia coli

<400> 340
Met Asn Leu Asn Ala Thr Ile Leu Gly Gln Ala Ile Ala Phe Val Leu
1 5 10 15
Phe Val Leu Phe Cys Met Lys Tyr Val Trp Pro Pro Leu Met Ala Ala
20 25 30
Ile Glu Lys Arg Gln Lys Glu Ile Ala Asp Gly Leu Ala Ser Ala Glu
35 40 45
Arg Ala His Lys Asp Leu Asp Leu Ala Lys Ala Ser Ala Thr Asp Gln
50 55 60
Leu Lys Lys Ala Lys Ala Glu Ala Gln Val Ile Ile Glu Gln Ala Asn
65 70 75 80
Lys Arg Arg Ser Gln Ile Leu Asp Glu Ala Lys Ala Glu Ala Glu Gln
85 90 95
Glu Arg Thr Lys Ile Val Ala Gln Ala Gln Ala Glu Ile Glu Ala Glu
100 105 110
Arg Lys Arg Ala Arg Glu Glu Leu Arg Lys Gln Val Ala Ile Leu Ala
115 120 125
Val Ala Gly Ala Glu Lys Ile Ile Glu Arg Ser Val Asp Glu Ala Ala
130 135 140

Asn Ser Asp Ile Val Asp Lys Leu Val Ala Glu Leu
145 150 155

<210> 341
<211> 79
<212> PRT
<213> Escherichia coli

<400> 341
Met Glu Asn Leu Asn Met Asp Leu Leu Tyr Met Ala Ala Ala Val Met
1 5 10 15
Met Gly Leu Ala Ala Ile Gly Ala Ala Ile Gly Ile Gly Ile Leu Gly
20 25 30
Gly Lys Phe Leu Glu Gly Ala Ala Arg Gln Pro Asp Leu Ile Pro Leu
35 40 45
Leu Arg Thr Gln Phe Phe Ile Val Met Gly Leu Val Asp Ala Ile Pro
50 55 60
Met Ile Ala Val Gly Leu Gly Leu Tyr Val Met Phe Ala Val Ala
65 70 75

<210> 342
<211> 271
<212> PRT
<213> Escherichia coli

<400> 342
Met Ala Ser Glu Asn Met Thr Pro Gln Asp Tyr Ile Gly His His Leu
1 5 10 15
Asn Asn Leu Gln Leu Asp Leu Arg Thr Phe Ser Leu Val Asp Pro Gln
20 25 30
Asn Pro Pro Ala Thr Phe Trp Thr Ile Asn Ile Asp Ser Met Phe Phe
35 40 45
Ser Val Val Leu Gly Leu Leu Phe Leu Val Leu Phe Arg Ser Val Ala
50 55 60
Lys Lys Ala Thr Ser Gly Val Pro Gly Lys Phe Gln Thr Ala Ile Glu
65 70 75 80
Leu Val Ile Gly Phe Val Asn Gly Ser Val Lys Asp Met Tyr His Gly
85 90 95
Lys Ser Lys Leu Ile Ala Pro Leu Ala Leu Thr Ile Phe Val Trp Val
100 105 110
Phe Leu Met Asn Leu Met Asp Leu Leu Pro Ile Asp Leu Leu Pro Tyr
115 120 125
Ile Ala Glu His Val Leu Gly Leu Pro Ala Leu Arg Val Val Pro Ser
130 135 140
Ala Asp Val Asn Val Thr Leu Ser Met Ala Leu Gly Val Phe Ile Leu
145 150 155 160
Ile Leu Phe Tyr Ser Ile Lys Met Lys Gly Ile Gly Gly Phe Thr Lys
165 170 175
Glu Leu Thr Leu Gln Pro Phe Asn His Trp Ala Phe Ile Pro Val Asn
180 185 190
Leu Ile Leu Glu Gly Val Ser Leu Leu Ser Lys Pro Val Ser Leu Gly
195 200 205
Leu Arg Leu Phe Gly Asn Met Tyr Ala Gly Glu Leu Ile Phe Ile Leu
210 215 220
Ile Ala Gly Leu Leu Pro Trp Trp Ser Gln Trp Ile Leu Asn Val Pro
225 230 235 240
Trp Ala Ile Phe His Ile Leu Ile Ile Thr Leu Gln Ala Phe Ile Phe
245 250 255
Met Val Leu Thr Ile Val Tyr Leu Ser Met Ala Ser Glu Glu His

260 265 270

<210> 343
<211> 130
<212> PRT
<213> Escherichia coli

<400> 343
Met Lys Asn Val Met Ser Val Ser Leu Val Ser Arg Asn Val Ala Arg
1 5 10 15
Lys Leu Leu Leu Val Gln Leu Leu Val Val Ile Ala Ser Gly Leu Leu
20 25 30
Phe Ser Leu Lys Asp Pro Phe Trp Gly Val Ser Ala Ile Ser Gly Gly
35 40 45
Leu Ala Val Phe Leu Pro Asn Val Leu Phe Met Ile Phe Ala Trp Arg
50 55 60
His Gln Ala His Thr Pro Ala Lys Gly Arg Val Ala Trp Thr Phe Ala
65 70 75 80
Phe Gly Glu Ala Phe Lys Val Leu Ala Met Leu Val Leu Leu Val Val
85 90 95
Ala Leu Ala Val Leu Lys Ala Val Phe Leu Pro Leu Ile Val Thr Trp
100 105 110
Val Leu Val Leu Val Val Gln Ile Leu Ala Pro Ala Val Ile Asn Asn
115 120 125
Lys Gly
130

<210> 344
<211> 413
<212> PRT
<213> Escherichia coli

<400> 344
Met Gln Lys His Gly Asp Arg Tyr Val Trp Ile Asn Pro Pro Ala Ile
1 5 10 15
Pro Leu Ser Thr Glu Glu Met Asp Ser Val Phe Ala Leu Pro Tyr Lys
20 25 30
Arg Val Pro His Pro Ala Tyr Gly Asn Ala Arg Ile Pro Ala Tyr Glu
35 40 45
Met Ile Arg Phe Ser Val Asn Ile Met Arg Gly Cys Phe Gly Gly Cys
50 55 60
Ser Phe Cys Ser Ile Thr Glu His Glu Gly Arg Ile Ile Gln Ser Arg
65 70 75 80
Ser Glu Asp Ser Ile Ile Asn Glu Ile Glu Ala Ile Arg Asp Thr Val
85 90 95
Pro Gly Phe Thr Gly Val Ile Ser Asp Leu Gly Gly Pro Thr Ala Asn
100 105 110
Met Tyr Met Leu Arg Cys Lys Ser Pro Arg Ala Glu Gln Thr Cys Arg
115 120 125
Arg Leu Ser Cys Val Tyr Pro Asp Ile Cys Pro His Met Asp Thr Asn
130 135 140
His Glu Pro Thr Ile Asn Leu Tyr Arg Arg Ala Arg Asp Leu Lys Gly
145 150 155 160
Ile Lys Lys Ile Leu Ile Ala Ser Gly Val Arg Tyr Asp Ile Ala Val
165 170 175
Glu Asp Pro Arg Tyr Ile Lys Glu Leu Ala Thr His His Val Gly Gly
180 185 190
Tyr Leu Lys Ile Ala Pro Glu His Thr Glu Glu Gly Pro Leu Ser Lys
195 200 205

Met Met Lys Pro Gly Met Gly Ser Tyr Asp Arg Phe Lys Glu Leu Phe
    210   215   220
Asp Thr Tyr Ser Lys Gln Ala Gly Lys Glu Gln Tyr Leu Ile Pro Tyr
225   230   235   240
Phe Ile Ser Ala His Pro Gly Thr Arg Asp Glu Asp Met Val Asn Leu
                245   250   255
Ala Leu Trp Leu Lys Lys His Arg Phe Arg Leu Asp Gln Val Gln Asn
            260   265   270
Phe Tyr Pro Ser Pro Leu Ala Asn Ser Thr Thr Met Tyr Tyr Thr Gly
    275   280   285
Lys Asn Pro Leu Ala Lys Ile Gly Tyr Lys Ser Glu Asp Val Phe Val
    290   295   300
Pro Lys Gly Asp Lys Gln Arg Arg Leu His Lys Ala Leu Leu Arg Tyr
305   310   315   320
His Asp Pro Ala Asn Trp Pro Leu Ile Arg Gln Ala Leu Glu Ala Met
                325   330   335
Gly Lys Lys His Leu Ile Gly Ser Arg Arg Asp Cys Leu Val Pro Ala
            340   345   350
Pro Thr Ile Glu Glu Met Arg Glu Ala Arg Arg Gln Asn Arg Asn Thr
    355   360   365
Arg Pro Ala Leu Thr Lys His Thr Pro Met Ala Thr Gln Arg Gln Thr
    370   375   380
Pro Ala Thr Ala Lys Lys Ala Ser Ser Thr Gln Ser Arg Pro Val Asn
385   390   395   400
Ala Gly Ala Lys Lys Arg Pro Lys Ala Ala Val Gly Arg
                405   410

<210> 345
<211> 325
<212> PRT
<213> Escherichia coli

<400> 345
Met Ser Ser Ile Ser Leu Ile Gln Pro Asp Arg Asp Leu Phe Ser Trp
1   5   10   15
Pro Gln Tyr Trp Ala Ala Cys Phe Gly Pro Ala Pro Phe Leu Pro Met
            20   25   30
Ser Arg Glu Glu Met Asp Gln Leu Gly Trp Asp Ser Cys Asp Ile Ile
        35   40   45
Leu Val Thr Gly Asp Ala Tyr Val Asp His Pro Ser Phe Gly Met Ala
    50   55   60
Ile Cys Gly Arg Met Leu Glu Ala Gln Gly Phe Arg Val Gly Ile Ile
65   70   75   80
Ala Gln Pro Asp Trp Ser Ser Lys Asp Asp Phe Met Arg Leu Gly Lys
                85   90   95
Pro Asn Leu Phe Phe Gly Val Thr Ala Gly Asn Met Asp Ser Met Ile
            100   105   110
Asn Arg Tyr Thr Ala Asp Arg Arg Leu Arg His Asp Asp Ala Tyr Thr
        115   120   125
Pro Asp Asn Val Ala Gly Lys Arg Pro Asp Arg Ala Thr Leu Val Tyr
    130   135   140
Thr Gln Arg Cys Lys Glu Ala Trp Lys Asp Val Pro Val Ile Leu Gly
145   150   155   160
Gly Ile Glu Ala Ser Leu Arg Arg Thr Ala His Tyr Asp Tyr Trp Ser
                165   170   175
Asp Thr Val Arg Arg Ser Val Leu Val Asp Ser Lys Ala Asp Met Leu
            180   185   190
Met Phe Gly Asn Gly Glu Arg Pro Leu Val Glu Val Ala His Arg Leu
        195   200   205
Ala Met Gly Glu Pro Ile Ser Glu Ile Arg Asp Val Arg Asn Thr Ala

    210   215   220
Ile Ile Val Lys Glu Ala Leu Pro Gly Trp Ser Gly Val Asp Ser Thr
225   230   235   240
Arg Leu Asp Thr Pro Gly Lys Ile Asp Pro Ile Pro His Pro Tyr Gly
                245   250   255
Glu Asp Leu Pro Cys Ala Asp Asn Lys Pro Val Ala Pro Lys Lys Gln
            260   265   270
Glu Ala Lys Ala Val Thr Val Gln Pro Pro Arg Pro Lys Pro Trp Glu
        275   280   285
Lys Thr Tyr Val Leu Leu Pro Ser Phe Glu Lys Val Lys Gly Asp Lys
    290   295   300
Val Leu Tyr Ala His Ala Ser Arg Ile Leu His His Glu Thr Asn Pro
305   310   315   320
Ala Val Pro Ala His
                325

<210> 346
<211> 226
<212> PRT
<213> Escherichia coli

<400> 346
Met Ile Gln Tyr Leu Asn Val Phe Phe Tyr Asp Ile Tyr Pro Tyr Ile
 1   5   10   15
Cys Ala Thr Val Phe Phe Leu Gly Ser Trp Leu Arg Tyr Asp Tyr Gly
            20   25   30
Gln Tyr Thr Trp Arg Ala Ser Ser Ser Gln Met Leu Asp Lys Arg Gly
        35   40   45
Met Val Ile Trp Ser Asn Leu Phe His Ile Gly Ile Leu Gly Ile Phe
    50   55   60
Phe Gly His Leu Phe Gly Met Leu Thr Pro His Trp Met Tyr Ala Trp
65   70   75   80
Phe Leu Pro Val Ala Ala Lys Gln Leu Met Ala Met Val Leu Gly Gly
                85   90   95
Ile Cys Gly Val Leu Thr Leu Ile Gly Gly Ala Gly Leu Leu Trp Arg
            100   105   110
Arg Leu Thr Asn Gln Arg Val Arg Ala Thr Ser Thr Thr Pro Asp Ile
        115   120   125
Ile Ile Met Ser Ile Leu Leu Ile Gln Cys Leu Leu Gly Leu Ser Thr
    130   135   140
Ile Pro Phe Ser Ala Gln Tyr Pro Asp Gly Ser Glu Met Met Lys Leu
145   150   155   160
Val Gly Trp Ala Gln Ser Ile Val Thr Phe Arg Gly Gly Ser Ser Glu
                165   170   175
Met Leu Asn Gly Val Ala Phe Val Phe Arg Leu His Leu Val Leu Gly
            180   185   190
Met Thr Ile Phe Leu Leu Phe Pro Phe Thr Arg Leu Val His Val Trp
        195   200   205
Ser Ala Pro Phe Glu Tyr Phe Thr Arg Arg Tyr Gln Ile Val Arg Ser
    210   215   220
Arg Arg
225

<210> 347
<211> 231
<212> PRT
<213> Escherichia coli

<400> 347

Met Gln Ile Leu Lys Val Ile Gly Leu Leu Met Glu Tyr Pro Asp Glu
1 5 10 15
Leu Leu Trp Glu Cys Lys Glu Asp Ala Leu Ala Leu Ile Arg Arg Asp
20 25 30
Ala Pro Met Leu Thr Asp Phe Thr His Asn Leu Leu Asn Ala Pro Leu
35 40 45
Leu Asp Lys Gln Ala Glu Trp Cys Glu Val Phe Asp Arg Gly Arg Thr
50 55 60
Thr Ser Leu Leu Leu Phe Glu His Val His Ala Glu Ser Arg Asp Arg
65 70 75 80
Gly Gln Ala Met Val Asp Leu Leu Ala Glu Tyr Glu Lys Val Gly Leu
85 90 95
Gln Leu Asp Cys Arg Glu Leu Pro Asp Tyr Leu Pro Leu Tyr Leu Glu
100 105 110
Tyr Leu Ser Val Leu Pro Asp Asp Gln Ala Lys Glu Gly Leu Leu Asn
115 120 125
Val Ala Pro Ile Leu Ala Leu Leu Gly Gly Arg Leu Lys Gln Arg Glu
130 135 140
Ala Pro Trp Tyr Ala Leu Phe Asp Ala Leu Leu Gln Leu Ala Gly Ser
145 150 155 160
Ser Leu Ser Ser Asp Ser Val Thr Lys Gln Val Asn Ser Glu Glu Arg
165 170 175
Asp Asp Thr Arg Gln Ala Leu Asp Ala Val Trp Glu Glu Glu Gln Val
180 185 190
Lys Phe Ile Glu Asp Asn Ala Thr Ala Cys Asp Ser Ser Pro Leu Asn
195 200 205
Gln Tyr Gln Arg Arg Phe Ser Gln Asp Val Ala Pro Gln Tyr Val Asp
210 215 220
Ile Ser Ala Gly Gly Gly Lys
225 230

<210> 348
<211> 514
<212> PRT
<213> Escherichia coli

<400> 348
Met Lys Ile Arg Ser Gln Val Gly Met Val Leu Asn Leu Asp Lys Cys
1 5 10 15
Ile Gly Cys His Thr Cys Ser Val Thr Cys Lys Asn Val Trp Thr Gly
20 25 30
Arg Glu Gly Met Glu Tyr Ala Trp Phe Asn Asn Val Glu Thr Lys Pro
35 40 45
Gly Ile Gly Tyr Pro Lys Asn Trp Glu Asp Gln Glu Glu Trp Gln Gly
50 55 60
Gly Trp Val Arg Asp Val Asn Gly Lys Ile Arg Pro Arg Leu Gly Asn
65 70 75 80
Lys Met Gly Val Ile Thr Lys Ile Phe Ala Asn Pro Val Val Pro Gln
85 90 95
Ile Asp Asp Tyr Tyr Glu Pro Phe Thr Phe Asp Tyr Glu His Leu His
100 105 110
Ser Ala Pro Glu Gly Lys His Ile Pro Thr Ala Arg Pro Arg Ser Leu
115 120 125
Ile Asp Gly Lys Arg Met Asp Lys Val Ile Trp Gly Pro Asn Trp Glu
130 135 140
Glu Leu Leu Gly Gly Glu Phe Glu Lys Arg Ala Arg Asp Arg Asn Phe
145 150 155 160
Glu Ala Met Gln Lys Glu Met Tyr Gly Gln Phe Glu Asn Thr Phe Met
165 170 175
Met Tyr Leu Pro Arg Leu Cys Glu His Cys Leu Asn Pro Ser Cys Val

```
                180                 185                 190
Ala Thr Cys Pro Ser Gly Ala Ile Tyr Lys Arg Glu Glu Asp Gly Ile
            195                 200                 205
Val Leu Ile Asp Gln Asp Lys Cys Arg Gly Trp Arg Leu Cys Ile Ser
    210                 215                 220
Gly Cys Pro Tyr Lys Lys Ile Tyr Phe Asn Trp Lys Ser Gly Lys Ser
225                 230                 235                 240
Glu Lys Cys Ile Phe Cys Tyr Pro Arg Ile Glu Ser Gly Gln Pro Thr
                245                 250                 255
Val Cys Ser Glu Thr Cys Val Gly Arg Ile Arg Tyr Leu Gly Val Leu
            260                 265                 270
Leu Tyr Asp Ala Asp Arg Ile Glu Glu Ala Ala Ser Thr Glu Arg Glu
        275                 280                 285
Val Asp Leu Tyr Glu Arg Gln Cys Glu Val Phe Leu Asp Pro His Asp
    290                 295                 300
Pro Ser Val Ile Glu Glu Ala Leu Lys Gln Gly Ile Pro Gln Asn Val
305                 310                 315                 320
Ile Asp Ala Ala Gln Arg Ser Pro Val Tyr Lys Met Ala Met Asp Trp
                325                 330                 335
Lys Leu Ala Leu Pro Leu His Pro Glu Tyr Arg Thr Leu Pro Met Val
            340                 345                 350
Trp Tyr Val Pro Pro Leu Ser Pro Ile Gln Ser Tyr Ala Asp Ala Gly
        355                 360                 365
Gly Leu Pro Lys Ser Glu Gly Val Leu Pro Ala Ile Glu Ser Leu Arg
    370                 375                 380
Ile Pro Val Gln Tyr Leu Ala Asn Met Leu Ser Ala Gly Asp Thr Gly
385                 390                 395                 400
Pro Val Leu Arg Ala Leu Lys Arg Met Met Ala Met Arg His Tyr Met
                405                 410                 415
Arg Ser Gln Thr Val Glu Gly Val Thr Asp Thr Arg Ala Ile Asp Glu
            420                 425                 430
Val Gly Leu Ser Val Ala Gln Val Glu Glu Met Tyr Arg Tyr Leu Ala
        435                 440                 445
Ile Ala Asn Tyr Glu Asp Arg Phe Val Ile Pro Thr Ser His Arg Glu
    450                 455                 460
Met Ala Gly Asp Ala Phe Ala Glu Arg Asn Gly Cys Gly Phe Thr Phe
465                 470                 475                 480
Gly Asp Gly Cys His Gly Ser Asp Ser Lys Phe Asn Leu Phe Asn Ser
                485                 490                 495
Ser Arg Ile Asp Ala Ile Asn Ile Thr Glu Val Arg Asp Lys Ala Glu
            500                 505                 510
Gly Glu
```

<210> 349
<211> 1246
<212> PRT
<213> Escherichia coli

<400> 349

```
Met Ser Lys Leu Leu Asp Arg Phe Arg Tyr Phe Lys Gln Lys Gly Glu
 1               5                  10                  15
Thr Phe Ala Asp Gly His Gly Gln Val Met His Ser Asn Arg Asp Trp
            20                  25                  30
Glu Asp Ser Tyr Arg Gln Arg Trp Gln Phe Asp Lys Ile Val Arg Ser
        35                  40                  45
Thr His Gly Val Asn Cys Thr Gly Ser Cys Ser Trp Lys Ile Tyr Val
    50                  55                  60
Lys Asn Gly Leu Val Thr Trp Glu Ile Gln Gln Thr Asp Tyr Pro Arg
65                  70                  75                  80
```

Thr Arg Pro Asp Leu Pro Asn His Glu Pro Arg Gly Cys Pro Arg Gly
85 90 95
Ala Ser Tyr Ser Trp Tyr Leu Tyr Ser Ala Asn Arg Leu Lys Tyr Pro
100 105 110
Leu Ile Arg Lys Arg Leu Ile Glu Leu Trp Arg Glu Ala Leu Lys Gln
115 120 125
His Ser Asp Pro Val Leu Ala Trp Ala Ser Ile Met Asn Asp Pro Gln
130 135 140
Lys Cys Leu Ser Tyr Lys Gln Val Arg Gly Arg Gly Gly Phe Ile Arg
145 150 155 160
Ser Asn Trp Gln Glu Leu Asn Gln Leu Ile Ala Ala Ala Asn Val Trp
165 170 175
Thr Ile Lys Thr Tyr Gly Pro Asp Arg Val Ala Gly Phe Ser Pro Ile
180 185 190
Pro Ala Met Ser Met Val Ser Tyr Ala Ala Gly Thr Arg Tyr Leu Ser
195 200 205
Leu Leu Gly Gly Thr Cys Leu Ser Phe Tyr Asp Trp Tyr Cys Asp Leu
210 215 220
Pro Pro Ala Ser Pro Met Thr Trp Gly Glu Gln Thr Asp Val Pro Glu
225 230 235 240
Ser Ala Asp Trp Tyr Asn Ser Ser Tyr Ile Ile Ala Trp Gly Ser Asn
245 250 255
Val Pro Gln Thr Arg Thr Pro Asp Ala His Phe Phe Thr Glu Val Arg
260 265 270
Tyr Lys Gly Thr Lys Thr Ile Ala Ile Thr Pro Asp Tyr Ser Glu Val
275 280 285
Ala Lys Leu Cys Asp Gln Trp Leu Ala Pro Lys Gln Gly Thr Asp Ser
290 295 300
Ala Leu Ala Met Ala Met Gly His Val Ile Leu Lys Glu Phe His Leu
305 310 315 320
Asp Asn Pro Ser Asp Tyr Phe Ile Asn Tyr Cys Arg Arg Tyr Ser Asp
325 330 335
Met Pro Met Leu Val Met Leu Glu Pro Arg Asp Asp Gly Ser Tyr Val
340 345 350
Pro Gly Arg Met Ile Arg Ala Ser Asp Leu Val Asp Gly Leu Gly Glu
355 360 365
Ser Asn Asn Pro Gln Trp Lys Thr Val Ala Val Asn Thr Ala Gly Glu
370 375 380
Leu Val Val Pro Asn Gly Ser Ile Gly Phe Arg Trp Gly Glu Lys Gly
385 390 395 400
Lys Trp Asn Leu Glu Ser Ile Ala Ala Gly Thr Glu Thr Glu Leu Ser
405 410 415
Leu Thr Leu Leu Gly Gln His Asp Ala Val Ala Gly Val Ala Phe Pro
420 425 430
Tyr Phe Gly Gly Ile Glu Asn Pro His Phe Arg Ser Val Lys His Asn
435 440 445
Pro Val Leu Val Arg Gln Leu Pro Val Lys Asn Leu Thr Leu Val Asp
450 455 460
Gly Asn Thr Cys Pro Val Val Ser Val Tyr Asp Leu Val Leu Ala Asn
465 470 475 480
Tyr Gly Leu Asp Arg Gly Leu Glu Asp Glu Asn Ser Ala Lys Asp Tyr
485 490 495
Ala Glu Ile Lys Pro Tyr Thr Pro Ala Trp Gly Glu Gln Ile Thr Gly
500 505 510
Val Pro Arg Gln Tyr Ile Glu Thr Ile Ala Arg Glu Phe Ala Asp Thr
515 520 525
Ala His Lys Thr His Gly Arg Ser Met Ile Ile Leu Gly Ala Gly Val
530 535 540
Asn His Trp Tyr His Met Asp Met Asn Tyr Arg Gly Met Ile Asn Met
545 550 555 560
Leu Ile Phe Cys Gly Cys Val Gly Gln Ser Gly Gly Gly Trp Ala His

565 570 575
Tyr Val Gly Gln Glu Lys Leu Arg Pro Gln Thr Gly Trp Leu Pro Leu
580 585 590
Ala Phe Ala Leu Asp Trp Asn Arg Pro Pro Arg Gln Met Asn Ser Thr
595 600 605
Ser Phe Phe Tyr Asn His Ser Ser Gln Trp Arg Tyr Glu Lys Val Ser
610 615 620
Ala Gln Glu Leu Leu Ser Pro Leu Ala Asp Ala Ser Lys Tyr Ser Gly
625 630 635 640
His Leu Ile Asp Phe Asn Val Arg Ala Glu Arg Met Gly Trp Leu Pro
645 650 655
Ser Ala Pro Gln Leu Gly Arg Asn Pro Leu Gly Ile Lys Ala Glu Ala
660 665 670
Asp Lys Ala Gly Leu Ser Pro Thr Glu Phe Thr Ala Gln Ala Leu Lys
675 680 685
Ser Gly Asp Leu Arg Met Ala Cys Glu Gln Pro Asp Ser Ser Ser Asn
690 695 700
His Pro Arg Asn Leu Phe Val Trp Arg Ser Asn Leu Leu Gly Ser Ser
705 710 715 720
Gly Lys Gly His Glu Tyr Met Gln Lys Tyr Leu Leu Gly Thr Glu Ser
725 730 735
Gly Ile Gln Gly Glu Glu Leu Gly Ala Ser Asp Gly Ile Lys Pro Glu
740 745 750
Glu Val Glu Trp Gln Thr Ala Ala Ile Glu Gly Lys Leu Asp Leu Leu
755 760 765
Val Thr Leu Asp Phe Arg Met Ser Ser Thr Cys Leu Phe Ser Asp Ile
770 775 780
Val Leu Pro Thr Ala Thr Trp Tyr Glu Lys Asp Asp Met Asn Thr Ser
785 790 795 800
Asp Met His Pro Phe Ile His Pro Leu Ser Ala Ala Val Asp Pro Ala
805 810 815
Trp Glu Ser Arg Ser Asp Trp Glu Ile Tyr Lys Gly Ile Ala Lys Ala
820 825 830
Phe Ser Gln Val Cys Val Gly His Leu Gly Lys Glu Thr Asp Val Val
835 840 845
Leu Gln Pro Leu Leu His Asp Ser Pro Ala Glu Leu Ser Gln Pro Cys
850 855 860
Glu Val Leu Asp Trp Arg Lys Gly Glu Cys Asp Leu Ile Pro Gly Lys
865 870 875 880
Thr Ala Pro Asn Ile Val Ala Val Glu Arg Asp Tyr Pro Ala Thr Tyr
885 890 895
Glu Arg Phe Thr Ser Leu Gly Pro Leu Met Asp Lys Leu Gly Asn Gly
900 905 910
Gly Lys Gly Ile Ser Trp Asn Thr Gln Asp Glu Ile Asp Phe Leu Gly
915 920 925
Lys Leu Asn Tyr Thr Lys Arg Asp Gly Pro Ala Gln Gly Arg Pro Leu
930 935 940
Ile Asp Thr Ala Ile Asp Ala Ser Glu Val Ile Leu Ala Leu Ala Pro
945 950 955 960
Glu Thr Asn Gly His Val Ala Val Lys Ala Trp Gln Ala Leu Gly Glu
965 970 975
Ile Thr Gly Arg Glu His Thr His Leu Ala Leu His Lys Glu Asp Glu
980 985 990
Lys Ile Arg Phe Arg Asp Ile Gln Ala Gln Pro Arg Lys Ile Ile Ser
995 1000 1005
Ser Pro Thr Trp Ser Gly Leu Glu Ser Asp His Val Ser Tyr Asn Ala
1010 1015 1020
Gly Tyr Thr Asn Val His Glu Leu Ile Pro Trp Arg Thr Leu Ser Gly
1025 1030 1035 1040
Arg Gln Gln Leu Tyr Gln Asp His Pro Trp Met Arg Ala Phe Gly Glu
1045 1050 1055

Ser Leu Val Ala Tyr Arg Pro Pro Ile Asp Thr Arg Ser Val Ser Glu
1060 1065 1070
Met Arg Gln Ile Pro Pro Asn Gly Phe Pro Glu Lys Ala Leu Asn Phe
1075 1080 1085
Leu Thr Pro His Gln Lys Trp Gly Ile His Ser Thr Tyr Ser Glu Asn
1090 1095 1100
Leu Leu Met Leu Thr Leu Ser Arg Gly Gly Pro Ile Val Trp Ile Ser
1105 1110 1115 1120
Glu Thr Asp Ala Arg Glu Leu Thr Ile Val Asp Asn Asp Trp Val Glu
1125 1130 1135
Val Phe Asn Ala Asn Gly Ala Leu Thr Ala Arg Ala Val Val Ser Gln
1140 1145 1150
Arg Val Pro Pro Gly Met Thr Met Met Tyr His Ala Gln Glu Arg Ile
1155 1160 1165
Met Asn Ile Pro Gly Ser Glu Val Thr Gly Met Arg Gly Gly Ile His
1170 1175 1180
Asn Ser Val Thr Arg Val Cys Pro Lys Pro Thr His Met Ile Gly Gly
1185 1190 1195 1200
Tyr Ala Gln Leu Ala Trp Gly Phe Asn Tyr Tyr Gly Thr Val Gly Ser
1205 1210 1215
Asn Arg Asp Glu Phe Ile Met Ile Arg Lys Met Lys Asn Val Asn Trp
1220 1225 1230
Leu Asp Asp Glu Gly Arg Asp Gln Val Gln Glu Ala Lys Lys
1235 1240 1245

<210> 350
<211> 165
<212> PRT
<213> Escherichia coli

<400> 350
Met Asp Leu Ser Gln Leu Thr Pro Arg Arg Pro Tyr Leu Leu Arg Ala
1 5 10 15
Phe Tyr Glu Trp Leu Leu Asp Asn Gln Leu Thr Pro His Leu Val Val
20 25 30
Asp Val Thr Leu Pro Gly Val Gln Val Pro Met Glu Tyr Ala Arg Asp
35 40 45
Gly Gln Ile Val Leu Asn Ile Ala Pro Arg Ala Val Gly Asn Leu Glu
50 55 60
Leu Ala Asn Asp Glu Val Arg Phe Asn Ala Arg Phe Gly Gly Ile Pro
65 70 75 80
Arg Gln Val Ser Val Pro Leu Ala Ala Val Leu Ala Ile Tyr Ala Arg
85 90 95
Glu Asn Gly Ala Gly Thr Met Phe Glu Pro Glu Ala Ala Tyr Asp Glu
100 105 110
Asp Thr Ser Ile Met Asn Asp Glu Glu Ala Ser Ala Asp Asn Glu Thr
115 120 125
Val Met Ser Val Ile Asp Gly Asp Lys Pro Asp His Asp Asp Asp Thr
130 135 140
His Pro Asp Asp Glu Pro Pro Gln Pro Pro Arg Gly Gly Arg Pro Ala
145 150 155 160
Leu Arg Val Val Lys
165

<210> 351
<211> 212
<212> PRT
<213> Escherichia coli

<400> 351
Met Ala Val Ala Ala Asn Lys Arg Ser Val Met Thr Leu Phe Ser Gly
1 5 10 15
Pro Thr Asp Ile Tyr Ser His Gln Val Arg Ile Val Leu Ala Glu Lys
20 25 30
Gly Val Ser Phe Glu Ile Glu His Val Glu Lys Asp Asn Pro Pro Gln
35 40 45
Asp Leu Ile Asp Leu Asn Pro Asn Gln Ser Val Pro Thr Leu Val Asp
50 55 60
Arg Glu Leu Thr Leu Trp Glu Ser Arg Ile Ile Met Glu Tyr Leu Asp
65 70 75 80
Glu Arg Phe Pro His Pro Pro Leu Met Pro Val Tyr Pro Val Ala Arg
85 90 95
Gly Glu Ser Arg Leu Tyr Met His Arg Ile Glu Lys Asp Trp Tyr Thr
100 105 110
Leu Met Asn Thr Ile Ile Asn Gly Ser Ala Ser Glu Ala Asp Ala Ala
115 120 125
Arg Lys Gln Leu Arg Glu Glu Leu Leu Ala Ile Ala Pro Val Phe Gly
130 135 140
Gln Lys Pro Tyr Phe Leu Ser Asp Glu Phe Ser Leu Val Asp Cys Tyr
145 150 155 160
Leu Ala Pro Leu Leu Trp Arg Leu Pro Gln Leu Gly Ile Glu Phe Ser
165 170 175
Gly Pro Gly Ala Lys Glu Leu Lys Gly Tyr Met Thr Arg Val Phe Glu
180 185 190
Arg Asp Ser Phe Leu Ala Ser Leu Thr Glu Ala Glu Arg Glu Met Arg
195 200 205
Leu Gly Arg Ser
210

<210> 352
<211> 505
<212> PRT
<213> Escherichia coli

<400> 352
Met Ser Glu Gln His Ala Gln Gly Ala Asp Ala Val Val Asp Leu Asn
1 5 10 15
Asn Glu Leu Lys Thr Arg Arg Glu Lys Leu Ala Asn Leu Arg Glu Gln
20 25 30
Gly Ile Ala Phe Pro Asn Asp Phe Arg Arg Asp His Thr Ser Asp Gln
35 40 45
Leu His Ala Glu Phe Asp Gly Lys Glu Asn Glu Glu Leu Glu Ala Leu
50 55 60
Asn Ile Glu Val Ala Val Ala Gly Arg Met Met Thr Arg Arg Ile Met
65 70 75 80
Gly Lys Ala Ser Phe Val Thr Leu Gln Asp Val Gly Gly Arg Ile Gln
85 90 95
Leu Tyr Val Ala Arg Asp Asp Leu Pro Glu Gly Val Tyr Asn Glu Gln
100 105 110
Phe Lys Lys Trp Asp Leu Gly Asp Ile Leu Gly Ala Lys Gly Lys Leu
115 120 125
Phe Lys Thr Lys Thr Gly Glu Leu Ser Ile His Cys Thr Glu Leu Arg
130 135 140
Leu Leu Thr Lys Ala Leu Arg Pro Leu Pro Asp Lys Phe His Gly Leu
145 150 155 160
Gln Asp Gln Glu Ala Arg Tyr Arg Gln Arg Tyr Leu Asp Leu Ile Ser
165 170 175
Asn Asp Glu Ser Arg Asn Thr Phe Lys Val Arg Ser Gln Ile Leu Ser
180 185 190

Gly Ile Arg Gln Phe Met Val Asn Arg Gly Phe Met Glu Val Glu Thr
        195   200   205
Pro Met Met Gln Val Ile Pro Gly Gly Ala Ala Ala Arg Pro Phe Ile
    210   215   220
Thr His His Asn Ala Leu Asp Leu Asp Met Tyr Leu Arg Ile Ala Pro
225   230   235   240
Glu Leu Tyr Leu Lys Arg Leu Val Val Gly Gly Phe Glu Arg Val Phe
                245   250   255
Glu Ile Asn Arg Asn Phe Arg Asn Glu Gly Ile Ser Val Arg His Asn
            260   265   270
Pro Glu Phe Thr Met Met Glu Leu Tyr Met Ala Tyr Ala Asp Tyr Lys
        275   280   285
Asp Leu Ile Glu Leu Thr Glu Ser Leu Phe Arg Thr Leu Ala Gln Asp
    290   295   300
Ile Leu Gly Lys Thr Glu Val Thr Tyr Gly Asp Val Thr Leu Asp Phe
305   310   315   320
Gly Lys Pro Phe Glu Lys Leu Thr Met Arg Glu Ala Ile Lys Lys Tyr
                325   330   335
Arg Pro Glu Thr Asp Met Ala Asp Leu Asp Asn Phe Asp Ser Ala Lys
            340   345   350
Ala Ile Ala Glu Ser Ile Gly Ile His Val Glu Lys Ser Trp Gly Leu
        355   360   365
Gly Arg Ile Val Thr Glu Ile Phe Glu Glu Val Ala Glu Ala His Leu
    370   375   380
Ile Gln Pro Thr Phe Ile Thr Glu Tyr Pro Ala Glu Val Ser Pro Leu
385   390   395   400
Ala Arg Arg Asn Asp Val Asn Pro Glu Ile Thr Asp Arg Phe Glu Phe
                405   410   415
Phe Ile Gly Gly Arg Glu Ile Gly Asn Gly Phe Ser Glu Leu Asn Asp
            420   425   430
Ala Glu Asp Gln Ala Gln Arg Phe Leu Asp Gln Val Ala Ala Lys Asp
        435   440   445
Ala Gly Asp Asp Glu Ala Met Phe Tyr Asp Glu Asp Tyr Val Thr Ala
    450   455   460
Leu Glu His Gly Leu Pro Pro Thr Ala Gly Leu Gly Ile Gly Ile Asp
465   470   475   480
Arg Met Val Met Leu Phe Thr Asn Ser His Thr Ile Arg Asp Val Ile
                485   490   495
Leu Phe Pro Ala Met Arg Pro Val Lys
            500   505

<210> 353
<211> 365
<212> PRT
<213> Escherichia coli

<400> 353
Met Phe Glu Ile Asn Pro Val Asn Asn Arg Ile Gln Asp Leu Thr Glu
1   5   10   15
Arg Ser Asp Val Leu Arg Gly Tyr Leu Asp Tyr Asp Ala Lys Lys Glu
            20   25   30
Arg Leu Glu Glu Val Asn Ala Glu Leu Glu Gln Pro Asp Val Trp Asn
        35   40   45
Glu Pro Glu Arg Ala Gln Ala Leu Gly Lys Glu Arg Ser Ser Leu Glu
    50   55   60
Ala Val Val Asp Thr Leu Asp Gln Met Lys Gln Gly Leu Glu Asp Val
65   70   75   80
Ser Gly Leu Leu Glu Leu Ala Val Glu Ala Asp Asp Glu Glu Thr Phe
                85   90   95
Asn Glu Ala Val Ala Glu Leu Asp Ala Leu Glu Glu Lys Leu Ala Gln

```
            100                 105                 110
Leu Glu Phe Arg Arg Met Phe Ser Gly Glu Tyr Asp Ser Ala Asp Cys
        115                 120                 125
Tyr Leu Asp Ile Gln Ala Gly Ser Gly Gly Thr Glu Ala Gln Asp Trp
    130                 135                 140
Ala Ser Met Leu Glu Arg Met Tyr Leu Arg Trp Ala Glu Ser Arg Gly
145                 150                 155                 160
Phe Lys Thr Glu Ile Ile Glu Glu Ser Glu Gly Glu Val Ala Gly Ile
                165                 170                 175
Lys Ser Val Thr Ile Lys Ile Ser Gly Asp Tyr Ala Tyr Gly Trp Leu
            180                 185                 190
Arg Thr Glu Thr Gly Val His Arg Leu Val Arg Lys Ser Pro Phe Asp
        195                 200                 205
Ser Gly Gly Arg Arg His Thr Ser Phe Ser Ser Ala Phe Val Tyr Pro
    210                 215                 220
Glu Val Asp Asp Asp Ile Asp Ile Glu Ile Asn Pro Ala Asp Leu Arg
225                 230                 235                 240
Ile Asp Val Tyr Arg Thr Ser Gly Ala Gly Gly Gln His Val Asn Arg
                245                 250                 255
Thr Glu Ser Ala Val Arg Ile Thr His Ile Pro Thr Gly Ile Val Thr
            260                 265                 270
Gln Cys Gln Asn Asp Arg Ser Gln His Lys Asn Lys Asp Gln Ala Met
        275                 280                 285
Lys Gln Met Lys Ala Lys Leu Tyr Glu Leu Glu Met Gln Lys Lys Asn
    290                 295                 300
Ala Glu Lys Gln Ala Met Glu Asp Asn Lys Ser Asp Ile Gly Trp Gly
305                 310                 315                 320
Ser Gln Ile Arg Ser Tyr Val Leu Asp Asp Ser Arg Ile Lys Asp Leu
                325                 330                 335
Arg Thr Gly Val Glu Thr Arg Asn Thr Gln Ala Val Leu Asp Gly Ser
            340                 345                 350
Leu Asp Gln Phe Ile Glu Ala Ser Leu Lys Ala Gly Leu
        355                 360                 365
```

<210> 354
<211> 577
<212> PRT
<213> Escherichia coli

<400> 354

```
Met Lys Gln Gln Ile Gln Leu Arg Arg Arg Glu Val Asp Glu Thr Ala
 1               5                  10                  15
Asp Leu Pro Ala Glu Leu Pro Pro Leu Leu Arg Arg Leu Tyr Ala Ser
            20                  25                  30
Arg Gly Val Arg Ser Ala Gln Glu Leu Glu Arg Ser Val Lys Gly Met
        35                  40                  45
Leu Pro Trp Gln Gln Leu Ser Gly Val Glu Lys Ala Val Glu Ile Leu
    50                  55                  60
Tyr Asn Ala Phe Arg Glu Gly Thr Arg Ile Ile Val Val Gly Asp Phe
65                  70                  75                  80
Asp Ala Asp Gly Ala Thr Ser Thr Ala Leu Ser Val Leu Ala Met Arg
                85                  90                  95
Ser Leu Gly Cys Ser Asn Ile Asp Tyr Leu Val Pro Asn Arg Phe Glu
            100                 105                 110
Asp Gly Tyr Gly Leu Ser Pro Glu Val Val Asp Gln Ala His Ala Arg
        115                 120                 125
Gly Ala Gln Leu Ile Val Thr Val Asp Asn Gly Ile Ser Ser His Ala
    130                 135                 140
Gly Val Glu His Ala Arg Ser Leu Gly Ile Pro Val Ile Val Thr Asp
145                 150                 155                 160
```

His His Leu Pro Gly Asp Thr Leu Pro Ala Ala Glu Ala Ile Ile Asn
165 170 175
Pro Asn Leu Arg Asp Cys Asn Phe Pro Ser Lys Ser Leu Ala Gly Val
180 185 190
Gly Val Ala Phe Tyr Leu Met Leu Ala Leu Arg Thr Phe Leu Arg Asp
195 200 205
Gln Gly Trp Phe Asp Glu Arg Asn Ile Ala Ile Pro Asn Leu Ala Glu
210 215 220
Leu Leu Asp Leu Val Ala Leu Gly Thr Val Ala Asp Val Val Pro Leu
225 230 235 240
Asp Ala Asn Asn Arg Ile Leu Thr Trp Gln Gly Met Ser Arg Ile Arg
245 250 255
Ala Gly Lys Cys Arg Pro Gly Ile Lys Ala Leu Leu Glu Val Ala Asn
260 265 270
Arg Asp Ala Gln Lys Leu Ala Ala Ser Asp Leu Gly Phe Ala Leu Gly
275 280 285
Pro Arg Leu Asn Ala Ala Gly Arg Leu Asp Asp Met Ser Val Gly Val
290 295 300
Ala Leu Leu Leu Cys Asp Asn Ile Gly Glu Ala Arg Val Leu Ala Asn
305 310 315 320
Glu Leu Asp Ala Leu Asn Gln Thr Arg Lys Glu Ile Glu Gln Gly Met
325 330 335
Gln Ile Glu Ala Leu Thr Leu Cys Glu Lys Leu Glu Arg Ser Arg Asp
340 345 350
Thr Leu Pro Gly Gly Leu Ala Met Tyr His Pro Glu Trp His Gln Gly
355 360 365
Val Val Gly Ile Leu Ala Ser Arg Ile Lys Glu Arg Phe His Arg Pro
370 375 380
Val Ile Ala Phe Ala Pro Ala Gly Asp Gly Thr Leu Lys Gly Ser Gly
385 390 395 400
Arg Ser Ile Gln Gly Leu His Met Arg Asp Ala Leu Glu Arg Leu Asp
405 410 415
Thr Leu Tyr Pro Gly Met Met Leu Lys Phe Gly Gly His Ala Met Ala
420 425 430
Ala Gly Leu Ser Leu Glu Glu Asp Lys Phe Lys Leu Phe Gln Gln Arg
435 440 445
Phe Gly Glu Leu Val Thr Glu Trp Leu Asp Pro Ser Leu Leu Gln Gly
450 455 460
Glu Val Val Ser Asp Gly Pro Leu Ser Pro Ala Glu Met Thr Met Glu
465 470 475 480
Val Ala Gln Leu Leu Arg Asp Ala Gly Pro Trp Gly Gln Met Phe Pro
485 490 495
Glu Pro Leu Phe Asp Gly His Phe Arg Leu Leu Gln Gln Arg Leu Val
500 505 510
Gly Glu Arg His Leu Lys Val Met Val Glu Pro Val Gly Gly Gly Pro
515 520 525
Leu Leu Asp Gly Ile Ala Phe Asn Val Asp Thr Ala Leu Trp Pro Asp
530 535 540
Asn Gly Val Arg Glu Val Gln Leu Ala Tyr Lys Leu Asp Ile Asn Glu
545 550 555 560
Phe Arg Gly Asn Arg Ser Leu Gln Ile Ile Ile Asp Asn Ile Trp Pro
565 570 575
Ile

<210> 355
<211> 236
<212> PRT
<213> Escherichia coli

<400> 355
Met Lys Lys Gly Phe Met Leu Phe Thr Leu Leu Ala Ala Phe Ser Gly
1 5 10 15
Phe Ala Gln Ala Asp Asp Ala Ala Ile Gln Gln Thr Leu Ala Lys Met
20 25 30
Gly Ile Lys Ser Ser Asp Ile Gln Pro Ala Pro Val Ala Gly Met Lys
35 40 45
Thr Val Leu Thr Asn Ser Gly Val Leu Tyr Ile Thr Asp Asp Gly Lys
50 55 60
His Ile Ile Gln Gly Pro Met Tyr Asp Val Ser Gly Thr Ala Pro Val
65 70 75 80
Asn Val Thr Asn Lys Met Leu Leu Lys Gln Leu Asn Ala Leu Glu Lys
85 90 95
Glu Met Ile Val Tyr Lys Ala Pro Gln Glu Lys His Val Ile Thr Val
100 105 110
Phe Thr Asp Ile Thr Cys Gly Tyr Cys His Lys Leu His Glu Gln Met
115 120 125
Ala Asp Tyr Asn Ala Leu Gly Ile Thr Val Arg Tyr Leu Ala Phe Pro
130 135 140
Arg Gln Gly Leu Asp Ser Asp Ala Glu Lys Glu Met Lys Ala Ile Trp
145 150 155 160
Cys Ala Lys Asp Lys Asn Lys Ala Phe Asp Asp Val Met Ala Gly Lys
165 170 175
Ser Val Ala Pro Ala Ser Cys Asp Val Asp Ile Ala Asp His Tyr Ala
180 185 190
Leu Gly Val Gln Leu Gly Val Ser Gly Thr Pro Ala Val Val Leu Ser
195 200 205
Asn Gly Thr Leu Val Pro Gly Tyr Gln Pro Pro Lys Glu Met Lys Glu
210 215 220
Phe Leu Asp Glu His Gln Lys Met Thr Ser Gly Lys
225 230 235

<210> 356
<211> 298
<212> PRT
<213> Escherichia coli

<400> 356
Met Lys Gln Asp Leu Ala Arg Ile Glu Gln Phe Leu Asp Ala Leu Trp
1 5 10 15
Leu Glu Lys Asn Leu Ala Glu Asn Thr Leu Asn Ala Tyr Arg Arg Asp
20 25 30
Leu Ser Met Met Val Glu Trp Leu His His Arg Gly Leu Thr Leu Ala
35 40 45
Thr Ala Gln Ser Asp Asp Leu Gln Ala Leu Leu Ala Glu Arg Leu Glu
50 55 60
Gly Gly Tyr Lys Ala Thr Ser Ser Ala Arg Leu Leu Ser Ala Val Arg
65 70 75 80
Arg Leu Phe Gln Tyr Leu Tyr Arg Glu Lys Phe Arg Glu Asp Asp Pro
85 90 95
Ser Ala His Leu Ala Ser Pro Lys Leu Pro Gln Arg Leu Pro Lys Asp
100 105 110
Leu Ser Glu Ala Gln Val Glu Arg Leu Leu Gln Ala Pro Leu Ile Asp
115 120 125
Gln Pro Leu Glu Leu Arg Asp Lys Ala Met Leu Glu Val Leu Tyr Ala
130 135 140
Thr Gly Leu Arg Val Ser Glu Leu Val Gly Leu Thr Met Ser Asp Ile
145 150 155 160
Ser Leu Arg Gln Gly Val Val Arg Val Ile Gly Lys Gly Asn Lys Glu
165 170 175

Arg Leu Val Pro Leu Gly Glu Glu Ala Val Tyr Trp Leu Glu Thr Tyr
180 185 190
Leu Glu His Gly Arg Pro Trp Leu Leu Asn Gly Val Ser Ile Asp Val
195 200 205
Leu Phe Pro Ser Gln Arg Ala Gln Gln Met Thr Arg Gln Thr Phe Trp
210 215 220
His Arg Ile Lys His Tyr Ala Val Leu Ala Gly Ile Asp Ser Glu Lys
225 230 235 240
Leu Ser Pro His Val Leu Arg His Ala Phe Ala Thr His Leu Leu Asn
245 250 255
His Gly Ala Asp Leu Arg Val Val Gln Met Leu Leu Gly His Ser Asp
260 265 270
Leu Ser Thr Thr Gln Ile Tyr Thr His Val Ala Thr Glu Arg Leu Arg
275 280 285
Gln Leu His Gln Gln His His Pro Arg Ala
290 295

<210> 357
<211> 367
<212> PRT
<213> Escherichia coli

<400> 357
Met Ser Asp Ser Gln Thr Leu Val Val Lys Leu Gly Thr Ser Val Leu
1 5 10 15
Thr Gly Gly Ser Arg Arg Leu Asn Arg Ala His Ile Val Glu Leu Val
20 25 30
Arg Gln Cys Ala Gln Leu His Ala Ala Gly His Arg Ile Val Ile Val
35 40 45
Thr Ser Gly Ala Ile Ala Ala Gly Arg Glu His Leu Gly Tyr Pro Glu
50 55 60
Leu Pro Ala Thr Ile Ala Ser Lys Gln Leu Leu Ala Ala Val Gly Gln
65 70 75 80
Ser Arg Leu Ile Gln Leu Trp Glu Gln Leu Phe Ser Ile Tyr Gly Ile
85 90 95
His Val Gly Gln Met Leu Leu Thr Arg Ala Asp Met Glu Asp Arg Glu
100 105 110
Arg Phe Leu Asn Ala Arg Asp Thr Leu Arg Ala Leu Leu Asp Asn Asn
115 120 125
Ile Val Pro Val Ile Asn Glu Asn Asp Ala Val Ala Thr Ala Glu Ile
130 135 140
Lys Val Gly Asp Asn Asp Asn Leu Ser Ala Leu Ala Ala Ile Leu Ala
145 150 155 160
Gly Ala Asp Lys Leu Leu Leu Leu Thr Asp Gln Lys Gly Leu Tyr Thr
165 170 175
Ala Asp Pro Arg Ser Asn Pro Gln Ala Glu Leu Ile Lys Asp Val Tyr
180 185 190
Gly Ile Asp Asp Ala Leu Arg Ala Ile Ala Gly Asp Ser Val Ser Gly
195 200 205
Leu Gly Thr Gly Gly Met Ser Thr Lys Leu Gln Ala Ala Asp Val Ala
210 215 220
Cys Arg Ala Gly Ile Asp Thr Ile Ile Ala Ala Gly Ser Lys Pro Gly
225 230 235 240
Val Ile Gly Asp Val Met Glu Gly Ile Ser Val Gly Thr Leu Phe His
245 250 255
Ala Gln Ala Thr Pro Leu Glu Asn Arg Lys Arg Trp Ile Phe Gly Ala
260 265 270
Pro Pro Ala Gly Glu Ile Thr Val Asp Glu Gly Ala Thr Ala Ala Ile
275 280 285
Leu Glu Arg Gly Ser Ser Leu Leu Pro Lys Gly Ile Lys Ser Val Thr

```
    290                 295                 300
Gly Asn Phe Ser Arg Gly Glu Val Ile Arg Ile Cys Asn Leu Glu Gly
305                 310                 315                 320
Arg Asp Ile Ala His Gly Val Ser Arg Tyr Asn Ser Asp Ala Leu Arg
                325                 330                 335
Arg Ile Ala Gly His His Ser Gln Glu Ile Asp Ala Ile Leu Gly Tyr
            340                 345                 350
Glu Tyr Gly Pro Val Ala Val His Arg Asp Asp Met Ile Thr Arg
        355                 360                 365


<210> 358
<211> 417
<212> PRT
<213> Escherichia coli

<400> 358
Met Leu Glu Gln Met Gly Ile Ala Ala Lys Gln Ala Ser Tyr Lys Leu
 1               5                  10                  15
Ala Gln Leu Ser Ser Arg Glu Lys Asn Arg Val Leu Glu Lys Ile Ala
            20                  25                  30
Asp Glu Leu Glu Ala Gln Ser Glu Ile Ile Leu Asn Ala Asn Ala Gln
        35                  40                  45
Asp Val Ala Asp Ala Arg Ala Asn Gly Leu Ser Glu Ala Met Leu Asp
    50                  55                  60
Arg Leu Ala Leu Thr Pro Ala Arg Leu Lys Gly Ile Ala Asp Asp Val
65                  70                  75                  80
Arg Gln Val Cys Asn Leu Ala Asp Pro Val Gly Gln Val Ile Asp Gly
                85                  90                  95
Gly Val Leu Asp Ser Gly Leu Arg Leu Glu Arg Arg Arg Val Pro Leu
            100                 105                 110
Gly Val Ile Gly Val Ile Tyr Glu Ala Arg Pro Asn Val Thr Val Asp
        115                 120                 125
Val Ala Ser Leu Cys Leu Lys Thr Gly Asn Ala Val Ile Leu Arg Gly
    130                 135                 140
Gly Lys Glu Thr Cys Arg Thr Asn Ala Ala Thr Val Ala Val Ile Gln
145                 150                 155                 160
Asp Ala Leu Lys Ser Cys Gly Leu Pro Ala Gly Ala Val Gln Ala Ile
                165                 170                 175
Asp Asn Pro Asp Arg Ala Leu Val Ser Glu Met Leu Arg Met Asp Lys
            180                 185                 190
Tyr Ile Asp Met Leu Ile Pro Arg Gly Gly Ala Gly Leu His Lys Leu
        195                 200                 205
Cys Arg Glu Gln Ser Thr Ile Pro Val Ile Thr Gly Gly Ile Gly Val
    210                 215                 220
Cys His Ile Tyr Val Asp Glu Ser Val Glu Ile Ala Glu Ala Leu Lys
225                 230                 235                 240
Val Ile Val Asn Ala Lys Thr Gln Arg Pro Ser Thr Cys Asn Thr Val
                245                 250                 255
Glu Thr Leu Leu Val Asn Lys Asn Ile Ala Asp Ser Phe Leu Pro Ala
            260                 265                 270
Leu Ser Lys Gln Met Ala Glu Ser Gly Val Thr Leu His Ala Asp Ala
        275                 280                 285
Ala Ala Leu Ala Gln Leu Gln Ala Gly Pro Ala Lys Val Val Ala Val
    290                 295                 300
Lys Ala Glu Glu Tyr Asp Asp Glu Phe Leu Ser Leu Asp Leu Asn Val
305                 310                 315                 320
Lys Ile Val Ser Asp Leu Asp Asp Ala Ile Ala His Ile Arg Glu His
                325                 330                 335
Gly Thr Gln His Ser Asp Ala Ile Leu Thr Arg Asp Met Arg Asn Ala
            340                 345                 350
```

Gln Arg Phe Val Asn Glu Val Asp Ser Ser Ala Val Tyr Val Asn Ala
355 360 365
Ser Thr Arg Phe Thr Asp Gly Gly Gln Phe Gly Leu Gly Ala Glu Val
370 375 380
Ala Val Ser Thr Gln Lys Leu His Ala Arg Gly Pro Met Gly Leu Glu
385 390 395 400
Ala Leu Thr Thr Tyr Lys Trp Ile Gly Ile Gly Asp Tyr Thr Ile Arg
405 410 415
Ala

<210> 359
<211> 186
<212> PRT
<213> Escherichia coli

<400> 359
Met Met Thr Arg Gln Ala Ser Met Lys Gly Phe Pro Ile Ala His Ile
1 5 10 15
Phe His Pro Ser Ile Pro Pro Met His Ala Val Val Asn Asn His Asn
20 25 30
Arg Asn Ile Asp Tyr Trp Thr Val Lys Arg Lys Phe Ala Glu Ile Val
35 40 45
Ser Thr Asn Asp Val Asn Lys Ile Tyr Ser Ile Ser Asn Glu Leu Arg
50 55 60
Arg Val Leu Ser Ala Ile Thr Ala Leu Asn Phe Tyr His Gly Asp Val
65 70 75 80
Pro Ser Val Met Ile Arg Ile Gln Pro Glu Asn Met Ser Pro Phe Ile
85 90 95
Ile Asp Ile Ser Thr Gly Glu His Asp Asp Tyr Ile Ile Gln Thr Leu
100 105 110
Asp Val Gly Thr Phe Ala Pro Phe Gly Glu Gln Cys Thr Cys Ser Ala
115 120 125
Val Asn Lys Lys Glu Leu Glu Cys Ile Lys Glu Thr Ile Ser Lys Tyr
130 135 140
Cys Ala Lys Phe Thr Arg Lys Glu Ala Ile Leu Thr Pro Leu Val His
145 150 155 160
Phe Asn Lys Thr Ser Ile Thr Ser Asp Cys Trp Gln Ile Leu Phe Phe
165 170 175
Ser Pro Asp His Phe Asn Asn Asp Phe Tyr
180 185

<210> 360
<211> 395
<212> PRT
<213> Escherichia coli

<400> 360
Met Phe Pro Val Ser Ser Ile Gly Asn Asp Ile Ser Ser Asp Leu Val
1 5 10 15
Arg Arg Lys Met Asn Asp Leu Pro Glu Ser Pro Thr Gly Asn Asn Leu
20 25 30
Glu Ala Leu Ala Pro Gly Ile Glu Lys Leu Lys Gln Thr Ser Ile Glu
35 40 45
Met Val Thr Leu Leu Asn Thr Leu Gln Pro Gly Gly Lys Cys Ile Ile
50 55 60
Thr Gly Asp Phe Gln Lys Glu Leu Ala Tyr Leu Gln Asn Val Ile Leu
65 70 75 80
Tyr Asn Val Ser Ser Leu Arg Leu Asp Phe Leu Gly Tyr Asn Ala Gln

```
                        85                  90                  95
Ile Ile Gln Arg Ser Asp Asn Thr Cys Glu Leu Thr Ile Asn Glu Pro
            100                 105                 110
Leu Lys Asn Gln Glu Ile Ser Thr Gly Asn Ile Asn Ile Asn Cys Pro
        115                 120                 125
Leu Lys Asp Ile Tyr Asn Glu Ile Arg Arg Leu Asn Val Ile Phe Ser
    130                 135                 140
Cys Gly Thr Gly Asp Ile Val Asp Leu Ser Ser Leu Asp Leu Arg Asn
145                 150                 155                 160
Val Asp Leu Asp Tyr Tyr Asp Phe Thr Asp Lys His Met Ala Asn Thr
                165                 170                 175
Ile Leu Asn Pro Phe Lys Leu Asn Ser Thr Asn Phe Thr Asn Ala Asn
            180                 185                 190
Met Phe Gln Val Asn Phe Val Ser Ser Thr Gln Asn Ala Thr Ile Ser
        195                 200                 205
Trp Asp Tyr Leu Leu Lys Ile Thr Pro Val Leu Ile Ser Ile Ser Asp
    210                 215                 220
Met Tyr Ser Glu Glu Lys Ile Lys Phe Val Glu Ser Cys Leu Asn Glu
225                 230                 235                 240
Pro Gly Asp Ile Thr Glu Glu Gln Leu Lys Ile Met Arg Phe Ala Ile
                245                 250                 255
Ile Lys Ser Ile Pro Arg Ala Thr Leu Thr Asp Lys Leu Glu Asn Glu
            260                 265                 270
Leu Thr Lys Glu Ile Tyr Lys Ser Ser Ser Lys Ile Ile Asn Cys Leu
        275                 280                 285
Asn Arg Ile Lys Leu Thr Glu Met Lys Glu Phe Ser Ser Glu Lys Ile
    290                 295                 300
Tyr Asp Tyr Ile Asp Ile Ile Ile Glu Asp Tyr Glu Asn Thr Lys Glu
305                 310                 315                 320
Asn Ala Tyr Leu Val Val Pro Gln Ile Asn Tyr Thr Met Asp Leu Asn
                325                 330                 335
Ile Glu Asp Ser Ser Ser Glu Glu Leu Leu Ser Asp Asn Thr Leu Glu
            340                 345                 350
Lys Asp Glu Asn Ser Pro Asp Asn Gly Phe Glu Val Gly Glu Tyr Asn
        355                 360                 365
Thr Tyr Glu Ala Tyr Asn Ser Glu Lys Gln Tyr Phe Thr Arg Glu Asp
    370                 375                 380
Tyr Thr Tyr Asp Tyr Asp Leu Leu Asn Ala Ile
385                 390                 395
```

<210> 361
<211> 315
<212> PRT
<213> Escherichia coli

<400> 361

```
Met Cys His Arg Ala Phe Arg Leu His Leu Cys Lys Asp Trp Val Phe
 1               5                  10                  15
Met Phe Ser Gly Leu Leu Ile Ile Leu Val Pro Leu Ile Val Gly Tyr
            20                  25                  30
Leu Ile Pro Leu Arg Gln Gln Ala Ala Leu Lys Val Ile Asn Gln Leu
        35                  40                  45
Leu Ser Trp Met Val Tyr Leu Ile Leu Phe Phe Met Gly Ile Ser Leu
    50                  55                  60
Ala Phe Leu Asp Asn Leu Ala Ser Asn Leu Leu Ala Ile Leu His Tyr
65                  70                  75                  80
Ser Ala Val Ser Ile Thr Val Ile Leu Leu Cys Asn Ile Ala Ala Leu
                85                  90                  95
Met Trp Leu Glu Arg Gly Leu Pro Trp Arg Asn His His Gln Gln Glu
            100                 105                 110
```

Lys Leu Pro Ser Arg Ile Ala Met Ala Leu Glu Ser Leu Lys Leu Cys
115 120 125
Gly Val Val Val Ile Gly Phe Ala Ile Gly Leu Ser Gly Leu Ala Phe
130 135 140
Leu Gln His Ala Thr Glu Ala Ser Glu Tyr Thr Leu Ile Leu Leu Leu
145 150 155 160
Phe Leu Val Gly Ile Gln Leu Arg Asn Asn Gly Met Thr Leu Lys Gln
165 170 175
Ile Val Leu Asn Arg Arg Gly Met Ile Val Ala Val Val Val Val Val
180 185 190
Ser Ser Leu Ile Gly Gly Leu Ile Asn Ala Phe Ile Leu Asp Leu Pro
195 200 205
Ile Asn Thr Ala Leu Ala Met Ala Ser Gly Phe Gly Trp Tyr Ser Leu
210 215 220
Ser Gly Ile Leu Leu Thr Glu Ser Phe Gly Pro Val Ile Gly Ser Ala
225 230 235 240
Ala Phe Phe Asn Asp Leu Ala Arg Glu Leu Ile Ala Ile Met Leu Ile
245 250 255
Pro Gly Leu Ile Arg Arg Ser Arg Ser Thr Ala Leu Gly Leu Cys Gly
260 265 270
Ala Thr Ser Met Asp Phe Thr Leu Pro Val Leu Gln Arg Thr Gly Gly
275 280 285
Leu Asp Met Val Pro Ala Ala Ile Val His Gly Phe Ile Leu Ser Leu
290 295 300
Leu Val Pro Ile Leu Ile Ala Phe Phe Ser Ala
305 310 315

<210> 362
<211> 96
<212> PRT
<213> Escherichia coli

<400> 362
Met Gln Asn Thr Thr His Asp Asn Val Ile Leu Glu Leu Thr Val Arg
1 5 10 15
Asn His Pro Gly Val Met Thr His Val Cys Gly Leu Phe Ala Arg Arg
20 25 30
Ala Phe Asn Val Glu Gly Ile Leu Cys Leu Pro Ile Gln Asp Ser Asp
35 40 45
Lys Ser His Ile Trp Leu Leu Val Asn Asp Asp Gln Arg Leu Glu Gln
50 55 60
Met Ile Ser Gln Ile Asp Lys Leu Glu Asp Val Val Lys Val Gln Arg
65 70 75 80
Asn Gln Ser Asp Pro Thr Met Phe Asn Lys Ile Ala Val Phe Phe Gln
85 90 95

<210> 363
<211> 562
<212> PRT
<213> Escherichia coli

<400> 363
Met Ala Ser Ser Gly Thr Thr Ser Thr Arg Lys Arg Phe Thr Gly Ala
1 5 10 15
Glu Phe Ile Val His Phe Leu Glu Gln Gln Gly Ile Lys Ile Val Thr
20 25 30
Gly Ile Pro Gly Gly Ser Ile Leu Pro Val Tyr Asp Ala Leu Ser Gln
35 40 45
Ser Thr Gln Ile Arg His Ile Leu Ala Arg His Glu Gln Gly Ala Gly

```
    50                  55                  60
Phe Ile Ala Gln Gly Met Ala Arg Thr Asp Gly Lys Pro Ala Val Cys
65                  70                  75                  80
Met Ala Cys Ser Gly Pro Gly Ala Thr Asn Leu Val Thr Ala Ile Ala
                85                  90                  95
Asp Ala Arg Leu Asp Ser Ile Pro Leu Ile Cys Ile Thr Gly Gln Val
            100                 105                 110
Pro Ala Ser Met Ile Gly Thr Asp Ala Phe Gln Glu Val Asp Thr Tyr
        115                 120                 125
Gly Ile Ser Ile Pro Ile Thr Lys His Asn Tyr Leu Val Arg His Ile
    130                 135                 140
Glu Glu Leu Pro Gln Val Met Ser Asp Ala Phe Arg Ile Ala Gln Ser
145                 150                 155                 160
Gly Arg Pro Gly Pro Val Trp Ile Asp Ile Pro Lys Asp Val Gln Thr
                165                 170                 175
Ala Val Phe Glu Ile Glu Thr Gln Pro Ala Met Ala Glu Lys Ala Ala
            180                 185                 190
Ala Pro Ala Phe Ser Glu Glu Ser Ile Arg Asp Ala Ala Ala Met Ile
        195                 200                 205
Asn Ala Ala Lys Arg Pro Val Leu Tyr Leu Gly Gly Gly Val Ile Asn
    210                 215                 220
Ala Pro Ala Arg Val Arg Glu Leu Ala Glu Lys Ala Gln Leu Pro Thr
225                 230                 235                 240
Thr Met Thr Leu Met Ala Leu Gly Met Leu Pro Lys Ala His Pro Leu
                245                 250                 255
Ser Leu Gly Met Leu Gly Met His Gly Val Arg Ser Thr Asn Tyr Ile
            260                 265                 270
Leu Gln Glu Ala Asp Leu Leu Ile Val Leu Gly Ala Arg Phe Asp Asp
        275                 280                 285
Arg Ala Ile Gly Lys Thr Glu Gln Phe Cys Pro Asn Ala Lys Ile Ile
    290                 295                 300
His Val Asp Ile Asp Arg Ala Glu Leu Gly Lys Ile Lys Gln Pro His
305                 310                 315                 320
Val Ala Ile Gln Ala Asp Val Asp Asp Val Leu Ala Gln Leu Ile Pro
                325                 330                 335
Leu Val Glu Ala Gln Pro Arg Ala Glu Trp His Gln Leu Val Ala Asp
            340                 345                 350
Leu Gln Arg Glu Phe Pro Cys Pro Ile Pro Lys Ala Cys Asp Pro Leu
        355                 360                 365
Ser His Tyr Gly Leu Ile Asn Ala Val Ala Ala Cys Val Asp Asp Asn
    370                 375                 380
Ala Ile Ile Thr Thr Asp Val Gly Gln His Gln Met Trp Thr Ala Gln
385                 390                 395                 400
Ala Tyr Pro Leu Asn Arg Pro Arg Gln Trp Leu Thr Ser Gly Gly Leu
                405                 410                 415
Gly Thr Met Gly Phe Gly Leu Pro Ala Ala Ile Gly Ala Ala Leu Ala
            420                 425                 430
Asn Pro Asp Arg Lys Val Leu Cys Phe Ser Gly Asp Gly Ser Leu Met
        435                 440                 445
Met Asn Ile Gln Glu Met Ala Thr Ala Ser Glu Asn Gln Leu Asp Val
    450                 455                 460
Lys Ile Ile Leu Met Asn Asn Glu Ala Leu Gly Leu Val His Gln Gln
465                 470                 475                 480
Gln Ser Leu Phe Tyr Glu Gln Gly Val Phe Ala Ala Thr Tyr Pro Gly
                485                 490                 495
Lys Ile Asn Phe Met Gln Ile Ala Ala Gly Phe Gly Leu Glu Thr Cys
            500                 505                 510
Asp Leu Asn Asn Glu Ala Asp Pro Gln Ala Ser Leu Gln Glu Ile Ile
        515                 520                 525
Asn Arg Pro Gly Pro Ala Leu Ile His Val Arg Ile Asp Ala Glu Glu
    530                 535                 540
```

Lys Val Tyr Pro Met Val Pro Pro Gly Ala Ala Asn Thr Glu Met Val
545 550 555 560
Gly Glu

<210> 364
<211> 32
<212> PRT
<213> Escherichia coli

<400> 364
Met Thr Thr Ser Met Leu Asn Ala Lys Leu Leu Pro Thr Ala Pro Ser
1 5 10 15
Ala Ala Val Val Val Val Arg Val Val Val Val Val Gly Asn Ala Pro
20 25 30

<210> 365
<211> 338
<212> PRT
<213> Escherichia coli

<400> 365
Met Phe Val Ile Trp Ser His Arg Thr Gly Phe Ile Met Ser His Gln
1 5 10 15
Leu Thr Phe Ala Asp Ser Glu Phe Ser Ser Lys Arg Arg Gln Thr Arg
20 25 30
Lys Glu Ile Phe Leu Ser Arg Met Glu Gln Ile Leu Pro Trp Gln Asn
35 40 45
Met Val Glu Val Ile Glu Pro Phe Tyr Pro Lys Ala Gly Asn Gly Arg
50 55 60
Arg Pro Tyr Pro Leu Glu Thr Met Leu Arg Ile His Cys Met Gln His
65 70 75 80
Trp Tyr Asn Leu Ser Asp Gly Ala Met Glu Asp Ala Leu Tyr Glu Ile
85 90 95
Ala Ser Met Arg Leu Phe Ala Arg Leu Ser Leu Asp Ser Ala Leu Pro
100 105 110
Asp Arg Thr Thr Ile Met Asn Phe Arg His Leu Leu Glu Gln His Gln
115 120 125
Leu Ala Arg Gln Leu Phe Lys Thr Ile Asn Arg Trp Leu Ala Glu Ala
130 135 140
Gly Val Met Met Thr Gln Gly Thr Leu Val Asp Ala Thr Ile Ile Glu
145 150 155 160
Ala Pro Ser Ser Thr Lys Asn Lys Glu Gln Gln Arg Asp Pro Glu Met
165 170 175
His Gln Thr Lys Lys Gly Asn Gln Trp His Phe Gly Met Lys Ala His
180 185 190
Ile Gly Val Asp Ala Lys Ser Gly Leu Thr His Ser Leu Val Thr Thr
195 200 205
Ala Ala Asn Glu His Asp Leu Asn Gln Leu Gly Asn Leu Leu His Gly
210 215 220
Glu Glu Gln Phe Val Ser Ala Asp Ala Gly Tyr Gln Gly Ala Pro Gln
225 230 235 240
Arg Glu Glu Leu Ala Glu Val Asp Val Asp Trp Leu Ile Ala Glu Arg
245 250 255
Pro Gly Lys Val Arg Thr Leu Lys Gln His Pro Arg Lys Asn Lys Thr
260 265 270
Ala Ile Asn Ile Glu Tyr Met Lys Ala Ser Ile Arg Ala Arg Val Glu
275 280 285
His Pro Phe Arg Ile Ile Lys Arg Gln Phe Gly Phe Val Lys Ala Arg

```
    290                 295                 300
Tyr Lys Gly Leu Leu Lys Asn Asp Asn Gln Leu Ala Met Leu Phe Thr
305                 310                 315                 320
Leu Ala Asn Leu Phe Arg Ala Asp Gln Met Ile Arg Gln Trp Glu Arg
                325                 330                 335
Ser His
```

<210> 366
<211> 157
<212> PRT
<213> Escherichia coli

<400> 366

```
Met Val Tyr Ile Ile Ile Val Ser His Gly His Glu Asp Tyr Ile Lys
 1               5                  10                  15
Lys Leu Leu Glu Asn Leu Asn Ala Asp Asp Glu His Tyr Lys Ile Ile
            20                  25                  30
Val Arg Asp Asn Lys Asp Ser Leu Leu Leu Lys Gln Ile Cys Gln His
        35                  40                  45
Tyr Ala Gly Leu Asp Tyr Ile Ser Gly Gly Val Tyr Gly Phe Gly His
    50                  55                  60
Asn Asn Asn Ile Ala Val Ala Tyr Val Lys Glu Lys Tyr Arg Pro Ala
65                  70                  75                  80
Asp Asp Asp Tyr Ile Leu Phe Leu Asn Pro Asp Ile Ile Met Lys His
                85                  90                  95
Asp Asp Leu Leu Thr Tyr Ile Lys Tyr Val Glu Ser Lys Arg Tyr Ala
            100                 105                 110
Phe Ser Thr Leu Cys Leu Phe Arg Asp Glu Ala Lys Ser Leu His Asp
        115                 120                 125
Tyr Ser Val Arg Lys Phe Pro Val Leu Ser Asp Phe Ile Val Ser Phe
    130                 135                 140
Met Leu Gly Ile Lys Glu Gly Ala Asn Lys Ser Leu Ile
145                 150                 155
```

<210> 367
<211> 372
<212> PRT
<213> Escherichia coli

<400> 367

```
Met Gly Lys Ser Ile Val Val Val Ser Ala Val Asn Phe Thr Thr Gly
 1               5                  10                  15
Gly Pro Phe Thr Ile Leu Lys Lys Phe Leu Ala Ala Thr Asn Asn Lys
            20                  25                  30
Glu Asn Val Ser Phe Ile Ala Leu Val His Ser Ala Lys Glu Leu Lys
        35                  40                  45
Glu Ser Tyr Pro Trp Val Lys Phe Ile Glu Phe Pro Glu Val Lys Gly
    50                  55                  60
Ser Trp Leu Lys Arg Leu His Phe Glu Tyr Val Val Cys Lys Lys Leu
65                  70                  75                  80
Ser Lys Glu Leu Asn Ala Thr His Trp Ile Cys Leu His Asp Ile Thr
                85                  90                  95
Ala Asn Val Val Thr Lys Lys Arg Tyr Val Tyr Cys His Asn Pro Ala
            100                 105                 110
Pro Phe Tyr Lys Gly Ile Leu Phe Arg Glu Ile Leu Met Glu Pro Ser
        115                 120                 125
Phe Phe Leu Phe Lys Met Leu Tyr Gly Leu Ile Tyr Lys Ile Asn Ile
    130                 135                 140
```

Lys Lys Asn Thr Ala Val Phe Val Gln Gln Phe Trp Met Lys Glu Lys
145 150 155 160
Phe Ile Lys Lys Tyr Ser Ile Asn Asn Ile Ile Val Ser Arg Pro Glu
165 170 175
Ile Lys Leu Ser Asp Lys Ser Gln Leu Thr Asp Asp Asp Ser Gln Phe
180 185 190
Lys Asn Asn Pro Ser Glu Leu Thr Ile Phe Tyr Pro Ala Val Pro Arg
195 200 205
Val Phe Lys Asn Tyr Glu Leu Ile Ile Ser Ala Ala Arg Lys Leu Lys
210 215 220
Glu Gln Ser Asn Ile Lys Phe Leu Leu Thr Ile Ser Gly Thr Glu Asn
225 230 235 240
Ala Tyr Ala Lys Tyr Ile Ile Ser Leu Ala Glu Gly Leu Asp Asn Val
245 250 255
His Phe Leu Gly Tyr Leu Asp Lys Glu Lys Ile Asp His Cys Tyr Asn
260 265 270
Ile Ser Asp Ile Val Cys Phe Pro Ser Arg Leu Glu Thr Trp Gly Leu
275 280 285
Pro Leu Ser Glu Ala Lys Glu Arg Gly Lys Trp Val Leu Ala Ser Asp
290 295 300
Phe Pro Phe Thr Arg Glu Thr Leu Gly Ser Tyr Glu Lys Lys Ala Phe
305 310 315 320
Phe Asp Ser Asn Asn Asp Asp Met Leu Val Lys Leu Ile Ile Asp Phe
325 330 335
Lys Lys Gly Asn Leu Lys Lys Asp Ile Ser Asp Ala Asn Phe Ile Tyr
340 345 350
Arg Asn Glu Asn Val Leu Val Gly Phe Asp Glu Leu Val Asn Phe Ile
355 360 365
Thr Glu Glu His
370

<210> 368
<211> 196
<212> PRT
<213> Escherichia coli

<400> 368
Met Ile Leu Lys Leu Ala Lys Arg Tyr Gly Leu Cys Gly Phe Ile Arg
1 5 10 15
Leu Val Arg Asp Val Leu Leu Thr Arg Val Phe Tyr Arg Asn Cys Arg
20 25 30
Ile Ile Arg Phe Pro Cys Tyr Ile Arg Asn Asp Gly Ser Ile Asn Phe
35 40 45
Gly Glu Asn Phe Thr Ser Gly Val Gly Leu Arg Leu Asp Ala Phe Gly
50 55 60
Arg Gly Val Ile Phe Phe Ser Asp Asn Val Gln Val Asn Asp Tyr Val
65 70 75 80
His Ile Ala Ser Ile Glu Ser Val Thr Ile Gly Arg Asp Thr Leu Ile
85 90 95
Ala Ser Lys Val Phe Ile Thr Asp His Asn His Gly Ser Phe Lys His
100 105 110
Ser Asp Pro Met Ser Ser Pro Asn Ile Pro Pro Asp Met Arg Thr Leu
115 120 125
Glu Ser Ser Ala Val Val Ile Gly Gln Arg Val Trp Leu Gly Glu Asn
130 135 140
Val Thr Val Leu Pro Gly Thr Ile Ile Gly Asn Gly Val Val Val Gly
145 150 155 160
Ala Asn Ser Val Val Arg Gly Ser Ile Pro Glu Asn Thr Val Ile Ala
165 170 175
Gly Val Pro Ala Lys Ile Ile Lys Lys Tyr Asn His Glu Thr Lys Leu

180 185 190
Trp Glu Lys Ala
195

<210> 369
<211> 330
<212> PRT
<213> Escherichia coli

<400> 369
Met Tyr Phe Leu Asn Asp Leu Asn Phe Ser Arg Arg Asp Ala Gly Phe
1 5 10 15
Lys Ala Arg Lys Asp Ala Leu Asp Ile Ala Ser Asp Tyr Glu Asn Ile
20 25 30
Ser Val Val Asn Ile Pro Leu Trp Gly Gly Val Val Gln Arg Ile Ile
35 40 45
Ser Ser Val Lys Leu Ser Thr Phe Leu Cys Gly Leu Glu Asn Lys Asp
50 55 60
Val Leu Ile Phe Asn Phe Pro Met Ala Lys Pro Phe Trp His Ile Leu
65 70 75 80
Ser Phe Phe His Arg Leu Leu Lys Phe Arg Ile Val Pro Leu Ile His
85 90 95
Asp Ile Asp Glu Leu Arg Gly Gly Gly Gly Ser Asp Ser Val Arg Leu
100 105 110
Ala Thr Cys Asp Met Val Ile Ser His Asn Pro Gln Met Thr Lys Tyr
115 120 125
Leu Ser Lys Tyr Met Ser Gln Asp Lys Ile Lys Asp Ile Lys Ile Phe
130 135 140
Asp Tyr Leu Val Ser Ser Asp Val Glu His Arg Asp Val Thr Asp Lys
145 150 155 160
Gln Arg Gly Val Ile Tyr Ala Gly Asn Leu Ser Arg His Lys Cys Ser
165 170 175
Phe Ile Tyr Thr Glu Gly Cys Asp Phe Thr Leu Phe Gly Val Asn Tyr
180 185 190
Glu Asn Lys Asp Asn Pro Lys Tyr Leu Gly Ser Phe Asp Ala Gln Ser
195 200 205
Pro Glu Lys Ile Asn Leu Pro Gly Met Gln Phe Gly Leu Ile Trp Asp
210 215 220
Gly Asp Ser Val Glu Thr Cys Ser Gly Ala Phe Gly Asp Tyr Leu Lys
225 230 235 240
Phe Asn Asn Pro His Lys Thr Ser Leu Tyr Leu Ser Met Glu Leu Pro
245 250 255
Val Phe Ile Trp Asp Lys Ala Ala Leu Ala Asp Phe Ile Val Asp Asn
260 265 270
Arg Ile Gly Tyr Ala Val Gly Ser Ile Lys Glu Met Gln Glu Ile Val
275 280 285
Asp Ser Met Thr Ile Glu Thr Tyr Lys Gln Ile Ser Glu Asn Thr Lys
290 295 300
Ile Ile Ser Gln Lys Ile Arg Thr Gly Ser Tyr Phe Arg Asp Val Leu
305 310 315 320
Glu Glu Val Ile Asp Asp Leu Lys Thr Arg
325 330

<210> 370
<211> 388
<212> PRT
<213> Escherichia coli

<400> 370

Met Ile Tyr Leu Val Ile Ser Val Phe Leu Ile Thr Ala Phe Ile Cys
1 5 10 15
Leu Tyr Leu Lys Lys Asp Ile Phe Tyr Pro Ala Val Cys Val Asn Ile
20 25 30
Ile Phe Ala Leu Val Leu Leu Gly Tyr Glu Ile Thr Ser Asp Ile Tyr
35 40 45
Ala Phe Gln Leu Asn Asp Ala Thr Leu Ile Phe Leu Leu Cys Asn Val
50 55 60
Leu Thr Phe Thr Leu Ser Cys Leu Leu Thr Glu Ser Val Leu Asp Leu
65 70 75 80
Asn Ile Arg Lys Val Asn Asn Ala Ile Tyr Ser Ile Pro Ser Lys Lys
85 90 95
Val His Asn Val Gly Leu Leu Val Ile Ser Phe Ser Met Ile Tyr Ile
100 105 110
Cys Met Arg Leu Ser Asn Tyr Gln Phe Gly Thr Ser Leu Leu Ser Tyr
115 120 125
Met Asn Leu Ile Arg Asp Ala Asp Val Glu Asp Thr Ser Arg Asn Phe
130 135 140
Ser Ala Tyr Met Gln Pro Ile Ile Leu Thr Thr Phe Ala Leu Phe Ile
145 150 155 160
Trp Ser Lys Lys Phe Thr Asn Thr Lys Val Ser Lys Thr Phe Thr Leu
165 170 175
Leu Val Phe Ile Val Phe Ile Phe Ala Ile Ile Leu Asn Thr Gly Lys
180 185 190
Gln Ile Val Phe Met Val Ile Ile Ser Tyr Ala Phe Ile Val Gly Val
195 200 205
Asn Arg Val Lys His Tyr Val Tyr Leu Ile Thr Ala Val Gly Val Leu
210 215 220
Phe Ser Leu Tyr Met Leu Phe Leu Arg Gly Leu Pro Gly Gly Met Ala
225 230 235 240
Tyr Tyr Leu Ser Met Tyr Leu Val Ser Pro Ile Ile Ala Phe Gln Glu
245 250 255
Phe Tyr Phe Gln Gln Val Ser Asn Ser Ala Ser Ser His Val Phe Trp
260 265 270
Phe Phe Glu Arg Leu Met Gly Leu Leu Thr Gly Gly Val Ser Met Ser
275 280 285
Leu His Lys Glu Phe Val Trp Val Gly Leu Pro Thr Asn Val Tyr Thr
290 295 300
Ala Phe Ser Asp Tyr Val Tyr Ile Ser Ala Glu Leu Ser Tyr Leu Met
305 310 315 320
Met Val Ile His Gly Cys Ile Ser Gly Val Leu Trp Arg Leu Ser Arg
325 330 335
Asn Tyr Ile Ser Val Lys Ile Phe Tyr Ser Tyr Phe Ile Tyr Thr Phe
340 345 350
Ser Phe Ile Phe Tyr His Glu Ser Phe Met Thr Asn Ile Ser Ser Trp
355 360 365
Ile Gln Ile Thr Leu Cys Ile Ile Val Phe Ser Gln Phe Leu Lys Ala
370 375 380
Gln Lys Ile Lys
385

<210> 371
<211> 367
<212> PRT
<213> Escherichia coli

<400> 371
Met Tyr Asp Tyr Ile Ile Val Gly Ser Gly Leu Phe Gly Ala Val Cys
1 5 10 15
Ala Asn Glu Leu Lys Lys Leu Asn Lys Lys Val Leu Val Ile Glu Lys

```
                20                  25                  30
Arg Asn His Ile Gly Gly Asn Ala Tyr Thr Glu Asp Cys Glu Gly Ile
        35                  40                  45
Gln Ile His Lys Tyr Gly Ala His Ile Phe His Thr Asn Asp Lys Tyr
    50                  55                  60
Ile Trp Asp Tyr Val Asn Asp Leu Val Glu Phe Asn Arg Phe Thr Asn
65                  70                  75                  80
Ser Pro Leu Ala Ile Tyr Lys Asp Lys Leu Phe Asn Leu Pro Phe Asn
                85                  90                  95
Met Asn Thr Phe His Gln Met Trp Gly Val Lys Asp Pro Gln Glu Ala
            100                 105                 110
Gln Asn Ile Ile Asn Ala Gln Lys Lys Lys Tyr Gly Asp Lys Val Pro
        115                 120                 125
Glu Asn Leu Glu Glu Gln Ala Ile Ser Leu Val Gly Glu Asp Leu Tyr
    130                 135                 140
Gln Ala Leu Ile Lys Gly Tyr Thr Glu Lys Gln Trp Gly Arg Ser Ala
145                 150                 155                 160
Lys Glu Leu Pro Ala Phe Ile Ile Lys Arg Ile Pro Val Arg Phe Thr
                165                 170                 175
Phe Asp Asn Asn Tyr Phe Ser Asp Arg Tyr Gln Gly Ile Pro Val Gly
            180                 185                 190
Gly Tyr Thr Lys Leu Ile Glu Lys Met Leu Glu Gly Val Asp Val Lys
        195                 200                 205
Leu Gly Ile Asp Phe Leu Lys Asp Lys Asp Ser Leu Ala Ser Lys Ala
    210                 215                 220
His Arg Ile Ile Tyr Thr Gly Pro Ile Asp Gln Tyr Phe Asp Tyr Arg
225                 230                 235                 240
Phe Gly Ala Leu Glu Tyr Arg Ser Leu Lys Phe Glu Thr Glu Arg His
                245                 250                 255
Glu Phe Pro Asn Phe Gln Gly Asn Ala Val Ile Asn Phe Thr Asp Ala
            260                 265                 270
Asn Val Pro Tyr Thr Arg Ile Ile Glu His Lys His Phe Asp Tyr Val
        275                 280                 285
Glu Thr Lys His Thr Val Val Thr Lys Glu Tyr Pro Leu Glu Trp Lys
    290                 295                 300
Val Gly Asp Glu Pro Tyr Tyr Pro Val Asn Asp Asn Lys Asn Met Glu
305                 310                 315                 320
Leu Phe Lys Lys Tyr Arg Glu Leu Ala Ser Arg Glu Asp Lys Val Ile
                325                 330                 335
Phe Gly Gly Arg Leu Ala Glu Tyr Lys Tyr Tyr Asp Met His Gln Val
            340                 345                 350
Ile Ser Ala Ala Leu Tyr Gln Val Lys Asn Ile Met Ser Thr Asp
        355                 360                 365
```

<210> 372
<211> 230
<212> PRT
<213> Escherichia coli

<400> 372

```
Met Thr Ser Ser Ile Thr Asn Glu Ile Met Gln Leu Tyr Thr Asp Arg
 1               5                  10                  15
Glu Val Leu Asn Met Gly Leu Cys Ser Arg Tyr Lys Ser Leu Thr Cys
            20                  25                  30
Asn Ser Cys Ser Met His Cys Gln Ile Met Pro Glu Glu Ser Pro Arg
        35                  40                  45
Leu Gln Tyr Cys Ala Asn Ser Cys Phe Cys Met Trp Pro Glu Glu Ser
    50                  55                  60
Ser Tyr Phe Asn Arg Gly Val Val Glu Gly Ile Leu Thr Lys Asn His
65                  70                  75                  80
```

Asn Ala Arg Leu Ser Gly Tyr Ile Phe Val Asp Phe Ser Val Ser Phe
85 90 95
Leu Arg Leu Phe Leu Glu Lys Asp Trp Ile Asp Tyr Leu Ala Ser Thr
100 105 110
Asp Met Gly Ile Val Leu Val Ser Asp Arg Asn Met Gln Ser Leu Ala
115 120 125
Asn Tyr Trp Arg Lys His Asn Ser Ala Ile Ser Ala Val Ile Tyr Asn
130 135 140
Asp Asp Gly Leu Asp Val Ala Asn Glu Lys Ile Arg Gln Leu Phe Ile
145 150 155 160
Gly Arg Tyr Leu Ser Phe Thr Gly Gly Asn Thr Leu Thr Gln Met Glu
165 170 175
Phe Thr Ile Met Gly Tyr Met Val Ser Gly Tyr Asn Pro Tyr Gln Ile
180 185 190
Ala Glu Val Leu Asp Met Asp Ile Arg Ser Ile Tyr Ala Tyr Lys Gln
195 200 205
Arg Ile Glu Lys Arg Met Gly Gly Lys Ile Asn Glu Leu Phe Ile Arg
210 215 220
Ser His Ser Val Gln His
225 230

<210> 373
<211> 391
<212> PRT
<213> Escherichia coli

<400> 373
Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
1 5 10 15
Thr Ala Leu Thr Glu Leu Tyr Gly Met Glu Asn Pro Ser Ser Gln Pro
20 25 30
Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Arg Val
35 40 45
Gln Asn Ala Ala Gly Asp Ile Val Ser Leu Arg Asp Val Ile Glu Ser
50 55 60
Asp Lys Ser Thr Leu Leu Gly Glu Ala Val Ala Lys Arg Phe Gly Glu
65 70 75 80
Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Ala Gln Pro Leu Ser Ile
85 90 95
Gln Val His Pro Asn Lys His Asn Ser Glu Ile Gly Phe Ala Lys Glu
100 105 110
Asn Ala Ala Gly Ile Pro Met Asp Ala Ala Glu Arg Asn Tyr Lys Asp
115 120 125
Pro Asn His Lys Pro Glu Leu Val Phe Ala Leu Thr Pro Phe Leu Ala
130 135 140
Met Asn Ala Phe Arg Glu Phe Ser Glu Ile Val Ser Leu Leu Gln Pro
145 150 155 160
Val Ala Gly Ala His Pro Ala Ile Ala His Phe Leu Gln Gln Pro Asp
165 170 175
Ala Glu Arg Leu Ser Glu Leu Phe Ala Ser Leu Leu Asn Met Gln Gly
180 185 190
Glu Glu Lys Ser Arg Ala Leu Ala Ile Leu Lys Ser Ala Leu Asp Ser
195 200 205
Gln Gln Gly Glu Pro Trp Gln Thr Ile Arg Leu Ile Ser Glu Phe Tyr
210 215 220
Pro Glu Asp Ser Gly Leu Phe Ser Pro Leu Leu Leu Asn Val Val Lys
225 230 235 240
Leu Asn Pro Gly Glu Ala Met Phe Leu Phe Ala Glu Thr Pro His Ala
245 250 255
Tyr Leu Gln Gly Val Ala Leu Glu Val Met Ala Asn Ser Asp Asn Val

```
            260                 265                 270
Leu Arg Ala Gly Leu Thr Pro Lys Tyr Ile Asp Ile Pro Glu Leu Val
        275                 280                 285
Ala Asn Val Lys Phe Glu Ala Lys Pro Ala Asn Gln Leu Leu Thr Gln
    290                 295                 300
Pro Val Lys Gln Gly Ala Glu Leu Asp Phe Pro Ile Pro Val Asp Asp
305                 310                 315                 320
Phe Ala Phe Ser Leu His Asp Leu Ser Asp Lys Glu Thr Thr Ile Ser
                325                 330                 335
Gln Gln Ser Ala Ala Ile Leu Phe Cys Val Glu Gly Asp Ala Thr Leu
            340                 345                 350
Trp Lys Gly Ser Gln Gln Leu Gln Leu Lys Pro Gly Glu Ser Ala Phe
        355                 360                 365
Ile Ala Ala Asn Glu Ser Pro Val Thr Val Lys Gly His Gly Arg Leu
    370                 375                 380
Ala Arg Val Tyr Asn Lys Leu
385                 390
```

<210> 374
<211> 264
<212> PRT
<213> Escherichia coli

<400> 374

```
Met Lys Gln Tyr Leu Glu Leu Met Gln Lys Val Leu Asp Glu Gly Thr
 1               5                  10                  15
Gln Lys Asn Asp Arg Thr Gly Thr Gly Thr Leu Ser Ile Phe Gly His
            20                  25                  30
Gln Met Arg Phe Asn Leu Gln Asp Gly Phe Pro Leu Val Thr Thr Lys
        35                  40                  45
Arg Cys His Leu Arg Ser Ile Ile His Glu Leu Leu Trp Phe Leu Gln
    50                  55                  60
Gly Asp Thr Asn Ile Ala Tyr Leu His Glu Asn Asn Val Thr Ile Trp
65                  70                  75                  80
Asp Glu Trp Ala Asp Glu Asn Gly Asp Leu Gly Pro Val Tyr Gly Lys
                85                  90                  95
Gln Trp Arg Ala Trp Pro Thr Pro Asp Gly Arg His Ile Asp Gln Ile
            100                 105                 110
Thr Thr Val Leu Asn Gln Leu Lys Asn Asp Pro Asp Ser Arg Arg Ile
        115                 120                 125
Ile Val Ser Ala Trp Asn Val Gly Glu Leu Asp Lys Met Ala Leu Ala
    130                 135                 140
Pro Cys His Ala Phe Phe Gln Phe Tyr Val Ala Asp Gly Lys Leu Ser
145                 150                 155                 160
Cys Gln Leu Tyr Gln Arg Ser Cys Asp Val Phe Leu Gly Leu Pro Phe
                165                 170                 175
Asn Ile Ala Ser Tyr Ala Leu Leu Val His Met Met Ala Gln Gln Cys
            180                 185                 190
Asp Leu Glu Val Gly Asp Phe Val Trp Thr Gly Gly Asp Thr His Leu
        195                 200                 205
Tyr Ser Asn His Met Asp Gln Thr His Leu Gln Leu Ser Arg Glu Pro
    210                 215                 220
Arg Pro Leu Pro Lys Leu Ile Ile Lys Arg Lys Pro Glu Ser Ile Phe
225                 230                 235                 240
Asp Tyr Arg Phe Glu Asp Phe Glu Ile Glu Gly Tyr Asp Pro His Pro
                245                 250                 255
Gly Ile Lys Ala Pro Val Ala Ile
            260
```

<210> 375
<211> 291
<212> PRT
<213> Escherichia coli

<400> 375
Met Thr Ser Ser Tyr Leu His Phe Pro Glu Phe Asp Pro Val Ile Phe
1 5 10 15
Ser Ile Gly Pro Val Ala Leu His Trp Tyr Gly Leu Met Tyr Leu Val
20 25 30
Gly Phe Ile Phe Ala Met Trp Leu Ala Thr Arg Arg Ala Asn Arg Pro
35 40 45
Gly Ser Gly Trp Thr Lys Asn Glu Val Glu Asn Leu Leu Tyr Ala Gly
50 55 60
Phe Leu Gly Val Phe Leu Gly Gly Arg Ile Gly Tyr Val Leu Phe Tyr
65 70 75 80
Asn Phe Pro Gln Phe Met Ala Asp Pro Leu Tyr Leu Phe Arg Val Trp
85 90 95
Asp Gly Gly Met Ser Phe His Gly Gly Leu Ile Gly Val Ile Val Val
100 105 110
Met Ile Ile Phe Ala Arg Arg Thr Lys Arg Ser Phe Phe Gln Val Ser
115 120 125
Asp Phe Ile Ala Pro Leu Ile Pro Phe Gly Leu Gly Ala Gly Arg Leu
130 135 140
Gly Asn Phe Ile Asn Gly Glu Leu Trp Gly Arg Val Asp Pro Asn Phe
145 150 155 160
Pro Phe Ala Met Leu Phe Pro Gly Ser Arg Thr Glu Asp Ile Leu Leu
165 170 175
Leu Gln Thr Asn Pro Gln Trp Gln Ser Ile Phe Asp Thr Tyr Gly Val
180 185 190
Leu Pro Arg His Pro Ser Gln Leu Tyr Glu Leu Leu Leu Glu Gly Val
195 200 205
Val Leu Phe Ile Ile Leu Asn Leu Tyr Ile Arg Lys Pro Arg Pro Met
210 215 220
Gly Ala Val Ser Gly Leu Phe Leu Ile Gly Tyr Gly Ala Phe Arg Ile
225 230 235 240
Ile Val Glu Phe Phe Arg Gln Pro Asp Ala Gln Phe Thr Gly Ala Trp
245 250 255
Val Gln Tyr Ile Ser Met Gly Gln Ile Leu Ser Ile Pro Met Ile Val
260 265 270
Ala Gly Val Ile Met Met Val Trp Ala Tyr Arg Arg Ser Pro Gln Gln
275 280 285
His Val Ser
290

<210> 376
<211> 241
<212> PRT
<213> Escherichia coli

<400> 376
Met Asp Ser Leu Asn Leu Asn Lys His Ile Ser Gly Gln Phe Asn Ala
1 5 10 15
Glu Leu Glu Ser Ile Arg Thr Gln Val Met Thr Met Gly Gly Met Val
20 25 30
Glu Gln Gln Leu Ser Asp Ala Ile Thr Ala Met His Asn Gln Asp Ser
35 40 45
Asp Leu Ala Lys Arg Val Ile Glu Gly Asp Lys Asn Val Asn Met Met
50 55 60
Glu Val Ala Ile Asp Glu Ala Cys Val Arg Ile Ile Ala Lys Arg Gln

```
65                  70                  75                  80
Pro Thr Ala Ser Asp Leu Arg Leu Val Met Val Ile Ser Lys Thr Ile
                85                  90                  95
Ala Glu Leu Glu Arg Ile Gly Asp Val Ala Asp Lys Ile Cys Arg Thr
            100                 105                 110
Ala Leu Glu Lys Phe Ser Gln Gln His Gln Pro Leu Leu Val Ser Leu
        115                 120                 125
Glu Ser Leu Gly Arg His Thr Ile Gln Met Leu His Asp Val Leu Asp
    130                 135                 140
Ala Phe Ala Arg Met Asp Ile Asp Glu Ala Val Arg Ile Tyr Arg Glu
145                 150                 155                 160
Asp Lys Lys Val Asp Gln Glu Tyr Glu Gly Ile Val Arg Gln Leu Met
                165                 170                 175
Thr Tyr Met Met Glu Asp Ser Arg Thr Ile Pro Ser Val Leu Thr Ala
            180                 185                 190
Leu Phe Cys Ala Arg Ser Ile Glu Arg Ile Gly Asp Arg Cys Gln Asn
        195                 200                 205
Ile Cys Glu Phe Ile Phe Tyr Tyr Val Lys Gly Gln Asp Phe Arg His
    210                 215                 220
Val Gly Gly Asp Glu Leu Asp Lys Leu Leu Ala Gly Lys Asp Ser Asp
225                 230                 235                 240
Lys
```

<210> 377
<211> 257
<212> PRT
<213> Escherichia coli

<400> 377

```
Met Ser Met Val Glu Thr Ala Pro Ser Lys Ile Gln Val Arg Asn Leu
 1               5                  10                  15
Asn Phe Tyr Tyr Gly Lys Phe His Ala Leu Lys Asn Ile Asn Leu Asp
            20                  25                  30
Ile Ala Lys Asn Gln Val Thr Ala Phe Ile Gly Pro Ser Gly Cys Gly
        35                  40                  45
Lys Ser Thr Leu Leu Arg Thr Phe Asn Lys Met Phe Glu Leu Tyr Pro
    50                  55                  60
Glu Gln Arg Ala Glu Gly Glu Ile Leu Leu Asp Gly Asp Asn Ile Leu
65                  70                  75                  80
Thr Asn Ser Gln Asp Ile Ala Leu Leu Arg Ala Lys Val Gly Met Val
                85                  90                  95
Phe Gln Lys Pro Thr Pro Phe Pro Met Ser Ile Tyr Asp Asn Ile Ala
            100                 105                 110
Phe Gly Val Arg Leu Phe Glu Lys Leu Ser Arg Ala Asp Met Asp Glu
        115                 120                 125
Arg Val Gln Trp Ala Leu Thr Lys Ala Ala Leu Trp Asn Glu Thr Lys
    130                 135                 140
Asp Lys Leu His Gln Ser Gly Tyr Ser Leu Ser Gly Gly Gln Gln Gln
145                 150                 155                 160
Arg Leu Cys Ile Ala Arg Gly Ile Ala Ile Arg Pro Glu Val Leu Leu
                165                 170                 175
Leu Asp Glu Pro Cys Ser Ala Leu Asp Pro Ile Ser Thr Gly Arg Ile
            180                 185                 190
Glu Glu Leu Ile Thr Glu Leu Lys Gln Asp Tyr Thr Val Val Ile Val
        195                 200                 205
Thr His Asn Met Gln Gln Ala Ala Arg Cys Ser Asp His Thr Ala Phe
    210                 215                 220
Met Tyr Leu Gly Glu Leu Ile Glu Phe Ser Asn Thr Asp Asp Leu Phe
225                 230                 235                 240
```

Thr Lys Pro Ala Lys Lys Gln Thr Glu Asp Tyr Ile Thr Gly Arg Tyr
245 250 255
Gly

<210> 378
<211> 296
<212> PRT
<213> Escherichia coli

<400> 378
Met Ala Met Val Glu Met Gln Thr Thr Ala Ala Leu Ala Glu Ser Arg
1 5 10 15
Arg Lys Met Gln Ala Arg Arg Arg Leu Lys Asn Arg Ile Ala Leu Thr
20 25 30
Leu Ser Met Ala Thr Met Ala Phe Gly Leu Phe Trp Leu Ile Trp Ile
35 40 45
Leu Met Ser Thr Ile Thr Arg Gly Ile Asp Gly Met Ser Leu Ala Leu
50 55 60
Phe Thr Glu Met Thr Pro Pro Pro Asn Thr Glu Gly Gly Gly Leu Ala
65 70 75 80
Asn Ala Leu Ala Gly Ser Gly Leu Leu Ile Leu Trp Ala Thr Val Phe
85 90 95
Gly Thr Pro Leu Gly Ile Met Ala Gly Ile Tyr Leu Ala Glu Tyr Gly
100 105 110
Arg Lys Ser Trp Leu Ala Glu Val Ile Arg Phe Ile Asn Asp Ile Leu
115 120 125
Leu Ser Ala Pro Ser Ile Val Val Gly Leu Phe Val Tyr Thr Ile Val
130 135 140
Val Ala Gln Met Glu His Phe Ser Gly Trp Ala Gly Val Ile Ala Leu
145 150 155 160
Ala Leu Leu Gln Val Pro Ile Val Ile Arg Thr Thr Glu Asn Met Leu
165 170 175
Lys Leu Val Pro Tyr Ser Leu Arg Glu Ala Ala Tyr Ala Leu Gly Thr
180 185 190
Pro Lys Trp Lys Met Ile Ser Ala Ile Thr Leu Lys Ala Ser Val Ser
195 200 205
Gly Ile Met Thr Gly Ile Leu Leu Ala Ile Ala Arg Ile Ala Gly Glu
210 215 220
Thr Ala Pro Leu Leu Phe Thr Ala Leu Ser Asn Gln Phe Trp Ser Thr
225 230 235 240
Asp Met Met Gln Pro Ile Ala Asn Leu Pro Val Thr Ile Phe Lys Phe
245 250 255
Ala Met Ser Pro Phe Ala Glu Trp Gln Gln Leu Ala Trp Ala Gly Val
260 265 270
Leu Ile Ile Thr Leu Cys Val Leu Leu Leu Asn Ile Leu Ala Arg Val
275 280 285
Val Phe Ala Lys Asn Lys His Gly
290 295

<210> 379
<211> 319
<212> PRT
<213> Escherichia coli

<400> 379
Met Ala Ala Thr Lys Pro Ala Phe Asn Pro Pro Gly Lys Lys Gly Asp
1 5 10 15
Ile Ile Phe Ser Val Leu Val Lys Leu Ala Ala Leu Ile Val Leu Leu

```
            20                  25                  30
Met Leu Gly Gly Ile Ile Val Ser Leu Ile Ile Ser Ser Trp Pro Ser
        35                  40                  45
Ile Gln Lys Phe Gly Leu Ala Phe Leu Trp Thr Lys Glu Trp Asp Ala
    50                  55                  60
Pro Asn Asp Ile Tyr Gly Ala Leu Val Pro Ile Tyr Gly Thr Leu Val
65                  70                  75                  80
Thr Ser Phe Ile Ala Leu Leu Ile Ala Val Pro Val Ser Phe Gly Ile
                85                  90                  95
Ala Leu Phe Leu Thr Glu Leu Ala Pro Gly Trp Leu Lys Arg Pro Leu
            100                 105                 110
Gly Ile Ala Ile Glu Leu Leu Ala Ala Ile Pro Ser Ile Val Tyr Gly
        115                 120                 125
Met Trp Gly Leu Phe Ile Phe Ala Pro Leu Phe Ala Val Tyr Phe Gln
    130                 135                 140
Glu Pro Val Gly Asn Ile Met Ser Asn Ile Pro Ile Val Gly Ala Leu
145                 150                 155                 160
Phe Ser Gly Pro Ala Phe Gly Ile Gly Ile Leu Ala Ala Gly Val Ile
                165                 170                 175
Leu Ala Ile Met Ile Ile Pro Tyr Ile Ala Ala Val Met Arg Asp Val
            180                 185                 190
Phe Glu Gln Thr Pro Val Met Met Lys Glu Ser Ala Tyr Gly Ile Gly
        195                 200                 205
Cys Thr Thr Trp Glu Val Ile Trp Arg Ile Val Leu Pro Phe Thr Lys
    210                 215                 220
Asn Gly Val Ile Gly Gly Ile Met Leu Gly Leu Gly Arg Ala Leu Gly
225                 230                 235                 240
Glu Thr Met Ala Val Thr Phe Ile Ile Gly Asn Thr Tyr Gln Leu Asp
                245                 250                 255
Ser Ala Ser Leu Tyr Met Pro Gly Asn Ser Ile Thr Ser Ala Leu Ala
            260                 265                 270
Asn Glu Phe Ala Glu Ala Glu Ser Gly Leu His Val Ala Ala Leu Met
        275                 280                 285
Glu Leu Gly Leu Ile Leu Phe Val Ile Thr Phe Ile Val Leu Ala Ala
    290                 295                 300
Ser Lys Phe Met Ile Met Arg Leu Ala Lys Asn Glu Gly Ala Arg
305                 310                 315
```

<210> 380
<211> 346
<212> PRT
<213> Escherichia coli

<400> 380

```
Met Lys Val Met Arg Thr Thr Val Ala Thr Val Val Ala Ala Thr Leu
 1               5                  10                  15
Ser Met Ser Ala Phe Ser Val Phe Ala Glu Ala Ser Leu Thr Gly Ala
            20                  25                  30
Gly Ala Thr Phe Pro Ala Pro Val Tyr Ala Lys Trp Ala Asp Thr Tyr
        35                  40                  45
Gln Lys Glu Thr Gly Asn Lys Val Asn Tyr Gln Gly Ile Gly Ser Ser
    50                  55                  60
Gly Gly Val Lys Gln Ile Ile Ala Asn Thr Val Asp Phe Gly Ala Ser
65                  70                  75                  80
Asp Ala Pro Leu Ser Asp Glu Lys Leu Ala Gln Glu Gly Leu Phe Gln
                85                  90                  95
Phe Pro Thr Val Ile Gly Gly Val Val Leu Ala Val Asn Ile Pro Gly
            100                 105                 110
Leu Lys Ser Gly Glu Leu Val Leu Asp Gly Lys Thr Leu Gly Asp Ile
        115                 120                 125
```

Tyr Leu Gly Lys Ile Lys Lys Trp Asp Asp Glu Ala Ile Ala Lys Leu
130 135 140
Asn Pro Gly Leu Lys Leu Pro Ser Gln Asn Ile Ala Val Val Arg Arg
145 150 155 160
Ala Asp Gly Ser Gly Thr Ser Phe Val Phe Thr Ser Tyr Leu Ala Lys
165 170 175
Val Asn Glu Glu Trp Lys Asn Asn Val Gly Thr Gly Ser Thr Val Lys
180 185 190
Trp Pro Ile Gly Leu Gly Gly Lys Gly Asn Asp Gly Ile Ala Ala Phe
195 200 205
Val Gln Arg Leu Pro Gly Ala Ile Gly Tyr Val Glu Tyr Ala Tyr Ala
210 215 220
Lys Gln Asn Asn Leu Ala Tyr Thr Lys Leu Ile Ser Ala Asp Gly Lys
225 230 235 240
Pro Val Ser Pro Thr Glu Glu Asn Phe Ala Asn Ala Ala Lys Gly Ala
245 250 255
Asp Trp Ser Lys Thr Phe Ala Gln Asp Leu Thr Asn Gln Lys Gly Glu
260 265 270
Asp Ala Trp Pro Ile Thr Ser Thr Thr Phe Ile Leu Ile His Lys Asp
275 280 285
Gln Lys Lys Pro Glu Gln Gly Thr Glu Val Leu Lys Phe Phe Asp Trp
290 295 300
Ala Tyr Lys Thr Gly Ala Lys Gln Ala Asn Asp Leu Asp Tyr Ala Ser
305 310 315 320
Leu Pro Asp Ser Val Val Glu Gln Val Arg Ala Ala Trp Lys Thr Asn
325 330 335
Ile Lys Asp Ser Ser Gly Lys Pro Leu Tyr
340 345

<210> 381
<211> 236
<212> PRT
<213> Escherichia coli

<400> 381
Met Gly Ser Gly Leu Val Asn Gly Gly Asp Tyr Phe Tyr Asn Asn Leu
1 5 10 15
Ser Phe Thr Val Thr Arg Tyr Asn Gly Ile Met Ala Thr Asp Ser Thr
20 25 30
Gln Cys Val Lys Lys Ser Arg Gly Arg Pro Lys Val Phe Asp Arg Asp
35 40 45
Ala Ala Leu Asp Lys Ala Met Lys Leu Phe Trp Gln His Gly Tyr Glu
50 55 60
Ala Thr Ser Leu Ala Asp Leu Val Glu Ala Thr Gly Ala Lys Ala Pro
65 70 75 80
Thr Leu Tyr Ala Glu Phe Thr Asn Lys Glu Gly Leu Phe Arg Ala Val
85 90 95
Leu Asp Arg Tyr Ile Asp Arg Phe Ala Ala Lys His Glu Ala Gln Leu
100 105 110
Phe Cys Glu Glu Lys Ser Val Glu Ser Ala Leu Ala Asp Tyr Phe Ala
115 120 125
Ala Ile Ala Asn Cys Phe Thr Ser Lys Asp Thr Pro Ala Gly Cys Phe
130 135 140
Met Ile Asn Asn Cys Thr Thr Leu Ser Pro Asp Ser Gly Asp Ile Ala
145 150 155 160
Asn Thr Leu Lys Ser Arg His Ala Met Gln Glu Arg Thr Leu Gln Gln
165 170 175
Phe Leu Cys Gln Arg Gln Ala Arg Gly Glu Ile Pro Pro His Cys Asp
180 185 190
Val Thr His Leu Ala Glu Phe Leu Asn Cys Ile Ile Gln Gly Met Ser

```
        195                 200                 205
Ile Ser Ala Arg Glu Gly Ala Ser Leu Glu Lys Leu Met Gln Ile Ala
    210                 215                 220
Gly Thr Thr Leu Arg Leu Trp Pro Glu Leu Val Lys
225                 230                 235


<210> 382
<211> 181
<212> PRT
<213> Escherichia coli

<400> 382
Met Gln Ala Lys Ile Ala Ala Ser Asn Thr Gly Glu Leu Asp Ala Leu
 1               5                  10                  15
Gln Gln Leu Gly Phe Ser Leu Val Glu Gly Glu Val Asp Leu Ala Leu
            20                  25                  30
Pro Val Asn Asn Ala Ser Asp Ser Gly Ala Val Val Ala Gln Glu Thr
        35                  40                  45
Asp Ile Pro Ala Leu Arg Gln Leu Ala Ser Ala Ala Phe Ala Gln Ser
    50                  55                  60
Arg Phe Arg Ala Pro Trp Tyr Ala Pro Asp Ala Ser Ser Arg Phe Tyr
65                  70                  75                  80
Ala Gln Trp Ile Glu Asn Ala Val Arg Gly Thr Phe Asp His Gln Cys
                85                  90                  95
Leu Ile Leu Arg Ala Ala Ser Gly Asp Ile Arg Gly Tyr Val Ser Leu
            100                 105                 110
Arg Glu Leu Asn Ala Thr Asp Ala Arg Ile Gly Leu Leu Ala Gly Arg
        115                 120                 125
Gly Ala Gly Ala Glu Leu Met Gln Thr Ala Leu Asn Trp Ala Tyr Arg
    130                 135                 140
Arg Gly Lys Thr Thr Leu Arg Val Ala Thr Gln Met Gly Asn Thr Ala
145                 150                 155                 160
Ala Leu Lys Arg Tyr Ile Gln Ser Gly Ala Asn Val Glu Ser Thr Ala
                165                 170                 175
Tyr Trp Leu Tyr Arg
            180


<210> 383
<211> 376
<212> PRT
<213> Escherichia coli

<400> 383
Met Ile Pro Phe Asn Ala Pro Pro Val Val Gly Thr Glu Leu Asp Tyr
 1               5                  10                  15
Met Gln Ser Ala Met Gly Ser Gly Lys Leu Cys Gly Asp Gly Gly Phe
            20                  25                  30
Thr Arg Arg Cys Gln Gln Trp Leu Glu Gln Arg Phe Gly Ser Ala Lys
        35                  40                  45
Val Leu Leu Thr Pro Ser Cys Thr Ala Ser Leu Glu Met Ala Ala Leu
    50                  55                  60
Leu Leu Asp Ile Gln Pro Gly Asp Glu Val Ile Met Pro Ser Tyr Thr
65                  70                  75                  80
Phe Val Ser Thr Ala Asn Ala Phe Val Leu Arg Gly Ala Lys Ile Val
                85                  90                  95
Phe Val Asp Val Arg Pro Asp Thr Met Asn Ile Asp Glu Thr Leu Ile
            100                 105                 110
Glu Ala Ala Ile Thr Asp Lys Thr Arg Val Ile Val Pro Val His Tyr
        115                 120                 125
```

Ala Gly Val Ala Cys Glu Met Asp Thr Ile Met Ala Leu Ala Lys Lys
    130   135   140
His Asn Leu Phe Val Val Glu Asp Ala Ala Gln Gly Val Met Ser Thr
145   150   155   160
Tyr Lys Gly Arg Ala Leu Gly Thr Ile Gly His Ile Gly Cys Phe Ser
                165   170   175
Phe His Glu Thr Lys Asn Tyr Thr Ala Gly Gly Glu Gly Gly Ala Thr
            180   185   190
Leu Ile Asn Asp Lys Ala Leu Ile Glu Arg Ala Glu Ile Ile Arg Glu
        195   200   205
Lys Gly Thr Asn Arg Ser Gln Phe Phe Arg Gly Gln Val Asp Lys Tyr
    210   215   220
Thr Trp Arg Asp Ile Gly Ser Ser Tyr Leu Met Ser Asp Leu Gln Ala
225   230   235   240
Ala Tyr Leu Trp Ala Gln Leu Glu Ala Ala Asp Arg Ile Asn Gln Gln
                245   250   255
Arg Leu Ala Leu Trp Gln Asn Tyr Tyr Asp Ala Leu Ala Pro Leu Ala
            260   265   270
Lys Ala Gly Arg Ile Glu Leu Pro Ser Ile Pro Asp Gly Cys Val Gln
        275   280   285
Asn Ala His Met Phe Tyr Ile Lys Leu Arg Asp Ile Asp Asp Arg Ser
    290   295   300
Ala Leu Ile Asn Phe Leu Lys Glu Ala Glu Ile Met Ala Val Phe His
305   310   315   320
Tyr Ile Pro Leu His Gly Cys Pro Ala Gly Glu His Phe Gly Glu Phe
                325   330   335
His Gly Glu Asp Arg Tyr Thr Thr Lys Glu Ser Glu Arg Leu Leu Arg
            340   345   350
Leu Pro Leu Phe Tyr Asn Leu Ser Pro Val Asn Gln Arg Thr Val Ile
        355   360   365
Ala Thr Leu Leu Asn Tyr Phe Ser
    370   375

<210> 384
<211> 416
<212> PRT
<213> Escherichia coli

<400> 384
Met Ser Leu Ala Lys Ala Ser Leu Trp Thr Ala Ala Ser Thr Leu Val
 1   5   10   15
Lys Ile Gly Ala Gly Leu Leu Val Gly Lys Leu Leu Ala Val Ser Phe
            20   25   30
Gly Pro Ala Gly Leu Gly Leu Ala Ala Asn Phe Arg Gln Leu Ile Thr
        35   40   45
Val Leu Gly Val Leu Ala Gly Ala Gly Ile Phe Asn Gly Val Thr Lys
    50   55   60
Tyr Val Ala Gln Tyr His Asp Asn Pro Gln Gln Leu Arg Arg Val Val
65   70   75   80
Gly Thr Ser Ser Ala Met Val Leu Gly Phe Ser Thr Leu Met Ala Leu
                85   90   95
Val Phe Val Leu Ala Ala Ala Pro Ile Ser Gln Gly Leu Phe Gly Asn
            100   105   110
Thr Asp Tyr Gln Gly Leu Val Arg Leu Val Ala Leu Val Gln Met Gly
        115   120   125
Ile Ala Trp Gly Asn Leu Leu Leu Ala Leu Met Lys Gly Phe Arg Asp
    130   135   140
Ala Ala Gly Asn Ala Leu Ser Leu Ile Val Gly Ser Leu Ile Gly Val
145   150   155   160
Leu Ala Tyr Tyr Val Ser Tyr Arg Leu Gly Gly Tyr Glu Gly Ala Leu

```
                165                 170                 175
Leu Gly Leu Ala Leu Ile Pro Ala Leu Val Val Ile Pro Ala Ala Ile
            180                 185                 190
Met Leu Ile Lys Arg Gly Val Ile Pro Leu Ser Tyr Leu Lys Pro Ser
        195                 200                 205
Trp Asp Asn Gly Leu Ala Gly Gln Leu Ser Lys Phe Thr Leu Met Ala
    210                 215                 220
Leu Ile Thr Ser Val Thr Leu Pro Val Ala Tyr Ile Met Met Arg Lys
225                 230                 235                 240
Leu Leu Ala Ala Gln Tyr Ser Trp Asp Glu Val Gly Ile Trp Gln Gly
                245                 250                 255
Val Ser Ser Ile Ser Asp Ala Tyr Leu Gln Phe Ile Thr Ala Ser Phe
            260                 265                 270
Ser Val Tyr Leu Leu Pro Thr Leu Ser Arg Leu Thr Glu Lys Arg Asp
        275                 280                 285
Ile Thr Arg Glu Val Val Lys Ser Leu Lys Phe Val Leu Pro Ala Val
    290                 295                 300
Ala Ala Ala Ser Phe Thr Val Trp Leu Leu Arg Asp Phe Ala Ile Trp
305                 310                 315                 320
Leu Leu Leu Ser Asn Lys Phe Thr Ala Met Arg Asp Leu Phe Ala Trp
                325                 330                 335
Gln Leu Val Gly Asp Val Leu Lys Val Gly Ala Tyr Val Phe Gly Tyr
            340                 345                 350
Leu Val Ile Ala Lys Ala Ser Leu Arg Phe Tyr Ile Leu Ala Glu Val
        355                 360                 365
Ser Gln Phe Thr Leu Leu Met Val Phe Ala His Trp Leu Ile Pro Ala
    370                 375                 380
His Gly Ala Leu Gly Ala Ala Gln Ala Tyr Met Ala Thr Tyr Ile Val
385                 390                 395                 400
Tyr Phe Ser Leu Cys Cys Gly Val Phe Leu Leu Trp Arg Arg Arg Ala
                405                 410                 415
```

<210> 385
<211> 450
<212> PRT
<213> Escherichia coli

<400> 385

```
Met Ser Leu Leu Gln Phe Ser Gly Leu Phe Val Val Trp Leu Leu Cys
 1               5                  10                  15
Thr Leu Phe Ile Ala Thr Leu Thr Trp Phe Glu Phe Arg Arg Val Arg
            20                  25                  30
Phe Asn Phe Asn Val Phe Phe Ser Leu Leu Phe Leu Leu Thr Phe Phe
        35                  40                  45
Phe Gly Phe Pro Leu Thr Ser Val Leu Val Phe Arg Phe Asp Val Gly
    50                  55                  60
Val Ala Pro Pro Glu Ile Leu Leu Gln Ala Leu Leu Ser Ala Gly Cys
65                  70                  75                  80
Phe Tyr Ala Val Tyr Tyr Val Thr Tyr Lys Thr Arg Leu Arg Lys Arg
                85                  90                  95
Val Ala Asp Val Pro Arg Arg Pro Leu Phe Thr Met Asn Arg Val Glu
            100                 105                 110
Thr Asn Leu Thr Trp Val Ile Leu Met Gly Ile Ala Leu Val Ser Val
        115                 120                 125
Gly Ile Phe Phe Met His Asn Gly Phe Leu Leu Phe Arg Leu Asn Ser
    130                 135                 140
Tyr Ser Gln Ile Phe Ser Ser Glu Val Ser Gly Val Ala Leu Lys Arg
145                 150                 155                 160
Phe Phe Tyr Phe Phe Ile Pro Ala Met Leu Val Val Tyr Phe Leu Arg
                165                 170                 175
```

Gln Asp Ser Lys Ala Trp Leu Phe Phe Leu Val Ser Thr Val Ala Phe
180 185 190
Gly Leu Leu Thr Tyr Met Ile Val Gly Gly Thr Arg Ala Asn Ile Ile
195 200 205
Ile Ala Phe Ala Ile Phe Leu Phe Ile Gly Ile Ile Arg Gly Trp Ile
210 215 220
Ser Leu Trp Met Leu Ala Ala Ala Gly Val Leu Gly Ile Val Gly Met
225 230 235 240
Phe Trp Leu Ala Leu Lys Arg Tyr Gly Met Asn Val Ser Gly Asp Glu
245 250 255
Ala Phe Tyr Thr Phe Leu Tyr Leu Thr Arg Asp Thr Phe Ser Pro Trp
260 265 270
Glu Asn Leu Ala Leu Leu Leu Gln Asn Tyr Asp Asn Ile Asp Phe Gln
275 280 285
Gly Leu Ala Pro Ile Val Arg Asp Phe Tyr Val Phe Ile Pro Ser Trp
290 295 300
Leu Trp Pro Gly Arg Pro Ser Met Val Leu Asn Ser Ala Asn Tyr Phe
305 310 315 320
Thr Trp Glu Val Leu Asn Asn His Ser Gly Leu Ala Ile Ser Pro Thr
325 330 335
Leu Ile Gly Ser Leu Val Val Met Gly Gly Ala Leu Phe Ile Pro Leu
340 345 350
Gly Ala Ile Val Val Gly Leu Ile Ile Lys Trp Phe Asp Trp Leu Tyr
355 360 365
Glu Leu Gly Asn Arg Glu Pro Asn Arg Tyr Lys Ala Ala Ile Leu His
370 375 380
Ser Phe Cys Phe Gly Ala Ile Phe Asn Met Ile Val Leu Ala Arg Glu
385 390 395 400
Gly Leu Asp Ser Phe Val Ser Arg Val Val Phe Phe Ile Val Val Phe
405 410 415
Gly Ala Cys Leu Met Ile Ala Lys Leu Leu Tyr Trp Leu Phe Glu Ser
420 425 430
Ala Gly Leu Ile His Lys Arg Thr Lys Ser Ser Leu Arg Thr Gln Val
435 440 445
Glu Gly
450

<210> 386
<211> 246
<212> PRT
<213> Escherichia coli

<400> 386
Met Asn Asn Asn Thr Thr Ala Pro Thr Tyr Thr Leu Arg Gly Leu Gln
1 5 10 15
Leu Ile Gly Trp Arg Asp Met Gln His Ala Leu Asp Tyr Leu Phe Ala
20 25 30
Asp Gly Gln Leu Lys Gln Gly Thr Leu Val Ala Ile Asn Ala Glu Lys
35 40 45
Met Leu Thr Ile Glu Asp Asn Ala Glu Val Arg Glu Leu Ile Asn Ala
50 55 60
Ala Glu Phe Lys Tyr Ala Asp Gly Ile Ser Val Val Arg Ser Val Arg
65 70 75 80
Lys Lys Tyr Pro Gln Ala Gln Val Ser Arg Val Ala Gly Ala Asp Leu
85 90 95
Trp Glu Glu Leu Met Ala Arg Ala Gly Lys Glu Gly Thr Pro Val Phe
100 105 110
Leu Val Gly Gly Lys Pro Glu Val Leu Ala Gln Thr Glu Ala Lys Leu
115 120 125
Arg Asn Gln Trp Asn Val Asn Ile Val Gly Ser Gln Asp Gly Tyr Phe

```
    130                 135                 140
Lys Pro Glu Gln Arg Gln Ala Leu Phe Glu Arg Ile His Ala Ser Gly
145                 150                 155                 160
Ala Gln Ile Val Thr Val Ala Met Gly Ser Pro Lys Gln Glu Ile Ile
                165                 170                 175
Met Arg Asp Cys Arg Leu Val His Pro Asp Ala Leu Tyr Met Gly Val
            180                 185                 190
Gly Gly Thr Tyr Asp Val Phe Thr Gly His Val Lys Arg Ala Pro Lys
        195                 200                 205
Ile Trp Gln Thr Leu Gly Leu Glu Trp Leu Tyr Arg Leu Leu Ser Gln
    210                 215                 220
Pro Ser Arg Ile Lys Arg Gln Leu Arg Leu Leu Arg Tyr Leu Arg Trp
225                 230                 235                 240
His Tyr Thr Gly Asn Leu
                245
```

<210> 387
<211> 74
<212> PRT
<213> Escherichia coli

<400> 387

```
Met Thr Val Leu Ile His Val Leu Gly Ser Asp Ile Pro His His Asn
 1               5                  10                  15
Arg Thr Val Leu Arg Phe Phe Asn Asp Ala Leu Ala Ala Thr Ser Glu
            20                  25                  30
His Ala Arg Glu Phe Met Val Val Gly Lys Asp Asp Gly Leu Ser Asp
        35                  40                  45
Ser Cys Pro Ala Leu Ser Val Gln Phe Phe Pro Trp Glu Lys Ile Ala
    50                  55                  60
Gly Gly Ser Gly His Arg Glu Ser Lys Ser
65                  70
```

<210> 388
<211> 204
<212> PRT
<213> Escherichia coli

<400> 388

```
Met Arg Gly Glu Leu Leu Phe Phe Pro Thr Arg Met Asp Pro Ser Leu
 1               5                  10                  15
Asn Thr Met Ala Asn Asp Arg Gln Arg Glu Gly Lys Met Thr Ile Leu
            20                  25                  30
Val Gly Asn Ser Gly Asp Arg Ser Asn Glu His Ile Ala Ala Leu Arg
        35                  40                  45
Ala Val His Gln Gln Phe Gly Asp Thr Val Lys Val Val Val Pro Met
    50                  55                  60
Gly Tyr Pro Pro Asn Asn Glu Ala Tyr Ile Glu Glu Val Arg Gln Ala
65                  70                  75                  80
Gly Leu Glu Leu Phe Ser Glu Glu Asn Leu Gln Ile Leu Ser Glu Lys
                85                  90                  95
Leu Glu Phe Asp Ala Tyr Leu Ala Leu Leu Arg Gln Cys Asp Leu Gly
            100                 105                 110
Tyr Phe Ile Phe Ala Arg Gln Gln Gly Ile Gly Thr Leu Cys Leu Leu
        115                 120                 125
Ile Gln Ala Gly Ile Pro Cys Val Leu Asn Arg Glu Asn Pro Phe Trp
    130                 135                 140
Gln Asp Met Thr Glu Gln His Leu Pro Val Leu Phe Thr Thr Asp Asp
145                 150                 155                 160
```

Leu Asn Glu Asp Ile Val Arg Glu Ala Gln Arg Gln Leu Ala Ser Val
165 170 175
Asp Lys Asn Thr Ile Ala Phe Phe Ser Pro Asn Tyr Leu Gln Gly Trp
180 185 190
Gln Arg Ala Leu Ala Ile Ala Ala Arg Glu Val Ala
195 200

<210> 389
<211> 182
<212> PRT
<213> Escherichia coli

<400> 389
Met Ile Arg Gln Arg Arg Arg Ala Leu Thr Pro Glu Gln Gln Gln Glu
1 5 10 15
Met Gly Gln Gln Ala Ala Thr Arg Met Met Thr Tyr Pro Pro Val Val
20 25 30
Met Ala His Thr Val Ala Val Phe Leu Ser Phe Asp Gly Glu Leu Asp
35 40 45
Thr Gln Pro Leu Ile Glu Gln Leu Trp Arg Ala Gly Lys Arg Val Tyr
50 55 60
Leu Pro Val Leu His Pro Phe Ser Ala Gly Asn Leu Leu Phe Leu Asn
65 70 75 80
Tyr His Pro Gln Ser Glu Leu Val Met Asn Arg Leu Lys Ile His Glu
85 90 95
Pro Lys Leu Asp Val Arg Asp Val Leu Pro Leu Ser Arg Leu Asp Val
100 105 110
Leu Ile Thr Pro Leu Val Ala Phe Asp Glu Tyr Gly Gln Arg Leu Gly
115 120 125
Met Gly Gly Gly Phe Tyr Asp Arg Thr Leu Gln Asn Trp Gln His Tyr
130 135 140
Lys Thr Gln Pro Val Gly Tyr Ala His Asp Cys Gln Leu Val Glu Lys
145 150 155 160
Leu Pro Val Glu Glu Trp Asp Ile Pro Leu Pro Ala Val Val Thr Pro
165 170 175
Ser Lys Val Trp Glu Trp
180

<210> 390
<211> 91
<212> PRT
<213> Escherichia coli

<400> 390
Met Ala Arg Val Thr Val Gln Asp Ala Val Glu Lys Ile Gly Asn Arg
1 5 10 15
Phe Asp Leu Val Leu Val Ala Ala Arg Arg Ala Arg Gln Met Gln Val
20 25 30
Gly Gly Lys Asp Pro Leu Val Pro Glu Glu Asn Asp Lys Thr Thr Val
35 40 45
Ile Ala Leu Arg Glu Ile Glu Glu Gly Leu Ile Asn Asn Gln Ile Leu
50 55 60
Asp Val Arg Glu Arg Gln Glu Gln Gln Glu Gln Glu Ala Ala Glu Leu
65 70 75 80
Gln Ala Val Thr Ala Ile Ala Glu Gly Arg Arg
85 90

<210> 391

<211> 702
<212> PRT
<213> Escherichia coli

<400> 391
Met Tyr Leu Phe Glu Ser Leu Asn Gln Leu Ile Gln Thr Tyr Leu Pro
1 5 10 15
Glu Asp Gln Ile Lys Arg Leu Arg Gln Ala Tyr Leu Val Ala Arg Asp
20 25 30
Ala His Glu Gly Gln Thr Arg Ser Ser Gly Glu Pro Tyr Ile Thr His
35 40 45
Pro Val Ala Val Ala Cys Ile Leu Ala Glu Met Lys Leu Asp Tyr Glu
50 55 60
Thr Leu Met Ala Ala Leu Leu His Asp Val Ile Glu Asp Thr Pro Ala
65 70 75 80
Thr Tyr Gln Asp Met Glu Gln Leu Phe Gly Lys Ser Val Ala Glu Leu
85 90 95
Val Glu Gly Val Ser Lys Leu Asp Lys Leu Lys Phe Arg Asp Lys Lys
100 105 110
Glu Ala Gln Ala Glu Asn Phe Arg Lys Met Ile Met Ala Met Val Gln
115 120 125
Asp Ile Arg Val Ile Leu Ile Lys Leu Ala Asp Arg Thr His Asn Met
130 135 140
Arg Thr Leu Gly Ser Leu Arg Pro Asp Lys Arg Arg Arg Ile Ala Arg
145 150 155 160
Glu Thr Leu Glu Ile Tyr Ser Pro Leu Ala His Arg Leu Gly Ile His
165 170 175
His Ile Lys Thr Glu Leu Glu Glu Leu Gly Phe Glu Ala Leu Tyr Pro
180 185 190
Asn Arg Tyr Arg Val Ile Lys Glu Val Val Lys Ala Ala Arg Gly Asn
195 200 205
Arg Lys Glu Met Ile Gln Lys Ile Leu Ser Glu Ile Glu Gly Arg Leu
210 215 220
Gln Glu Ala Gly Ile Pro Cys Arg Val Ser Gly Arg Glu Lys His Leu
225 230 235 240
Tyr Ser Ile Tyr Cys Lys Met Val Leu Lys Glu Gln Arg Phe His Ser
245 250 255
Ile Met Asp Ile Tyr Ala Phe Arg Val Ile Val Asn Asp Ser Asp Thr
260 265 270
Cys Tyr Arg Val Leu Gly Gln Met His Ser Leu Tyr Lys Pro Arg Pro
275 280 285
Gly Arg Val Lys Asp Tyr Ile Ala Ile Pro Lys Ala Asn Gly Tyr Gln
290 295 300
Ser Leu His Thr Ser Met Ile Gly Pro His Gly Val Pro Val Glu Val
305 310 315 320
Gln Ile Arg Thr Glu Asp Met Asp Gln Met Ala Glu Met Gly Val Ala
325 330 335
Ala His Trp Ala Tyr Lys Glu His Gly Glu Thr Ser Thr Thr Ala Gln
340 345 350
Ile Arg Ala Gln Arg Trp Met Gln Ser Leu Leu Glu Leu Gln Gln Ser
355 360 365
Ala Gly Ser Ser Phe Glu Phe Ile Glu Ser Val Lys Ser Asp Leu Phe
370 375 380
Pro Asp Glu Ile Tyr Val Phe Thr Pro Glu Gly Arg Ile Val Glu Leu
385 390 395 400
Pro Ala Gly Ala Thr Pro Val Asp Phe Ala Tyr Ala Val His Thr Asp
405 410 415
Ile Gly His Ala Cys Val Gly Ala Arg Val Asp Arg Gln Pro Tyr Pro
420 425 430
Leu Ser Gln Pro Leu Thr Ser Gly Gln Thr Val Glu Ile Ile Thr Ala
435 440 445

Pro Gly Ala Arg Pro Asn Ala Ala Trp Leu Asn Phe Val Val Ser Ser
450 455 460
Lys Ala Arg Ala Lys Ile Arg Gln Leu Leu Lys Asn Leu Lys Arg Asp
465 470 475 480
Asp Ser Val Ser Leu Gly Arg Arg Leu Leu Asn His Ala Leu Gly Gly
485 490 495
Ser Arg Lys Leu Asn Glu Ile Pro Gln Glu Asn Ile Gln Arg Glu Leu
500 505 510
Asp Arg Met Lys Leu Ala Thr Leu Asp Asp Leu Leu Ala Glu Ile Gly
515 520 525
Leu Gly Asn Ala Met Ser Val Val Val Ala Lys Asn Leu Gln His Gly
530 535 540
Asp Ala Ser Ile Pro Pro Ala Thr Gln Ser His Gly His Leu Pro Ile
545 550 555 560
Lys Gly Ala Asp Gly Val Leu Ile Thr Phe Ala Lys Cys Cys Arg Pro
565 570 575
Ile Pro Gly Asp Pro Ile Ile Ala His Val Ser Pro Gly Lys Gly Leu
580 585 590
Val Ile His His Glu Ser Cys Arg Asn Ile Arg Gly Tyr Gln Lys Glu
595 600 605
Pro Glu Lys Phe Met Ala Val Glu Trp Asp Lys Glu Thr Ala Gln Glu
610 615 620
Phe Ile Thr Glu Ile Lys Val Glu Met Phe Asn His Gln Gly Ala Leu
625 630 635 640
Ala Asn Leu Thr Ala Ala Ile Asn Thr Thr Thr Ser Asn Ile Gln Ser
645 650 655
Leu Asn Thr Glu Glu Lys Asp Gly Arg Val Tyr Ser Ala Phe Ile Arg
660 665 670
Leu Thr Ala Arg Asp Arg Val His Leu Ala Asn Ile Met Arg Lys Ile
675 680 685
Arg Val Met Pro Asp Val Ile Lys Val Thr Arg Asn Arg Asn
690 695 700

<210> 392
<211> 229
<212> PRT
<213> Escherichia coli

<400> 392
Met Asn Pro Thr Arg Tyr Ala Arg Ile Cys Glu Met Leu Ala Arg Arg
1 5 10 15
Gln Pro Asp Leu Thr Val Cys Met Glu Gln Val His Lys Pro His Asn
20 25 30
Val Ser Ala Ile Ile Arg Thr Ala Asp Ala Val Gly Val His Glu Val
35 40 45
His Ala Val Trp Pro Gly Ser Arg Met Arg Thr Met Ala Ser Ala Ala
50 55 60
Ala Gly Ser Asn Ser Trp Val Gln Val Lys Thr His Arg Thr Ile Gly
65 70 75 80
Asp Ala Val Ala His Leu Lys Gly Gln Gly Met Gln Ile Leu Ala Thr
85 90 95
His Leu Ser Asp Asn Ala Val Asp Phe Arg Glu Ile Asp Tyr Thr Arg
100 105 110
Pro Thr Cys Ile Leu Met Gly Gln Glu Lys Thr Gly Ile Thr Gln Glu
115 120 125
Ala Leu Ala Leu Ala Asp Gln Asp Ile Ile Ile Pro Met Ile Gly Met
130 135 140
Val Gln Ser Leu Asn Val Ser Val Ala Ser Ala Leu Ile Leu Tyr Glu
145 150 155 160
Ala Gln Arg Gln Arg Gln Asn Ala Gly Met Tyr Leu Arg Glu Asn Ser

165 170 175
Met Leu Pro Glu Ala Glu Gln Gln Arg Leu Leu Phe Glu Gly Gly Tyr
180 185 190
Pro Val Leu Ala Lys Val Ala Lys Arg Lys Gly Leu Pro Tyr Pro His
195 200 205
Val Asn Gln Gln Gly Glu Ile Glu Ala Asp Ala Asp Trp Trp Ala Thr
210 215 220
Met Gln Ala Ala Gly
225

<210> 393
<211> 693
<212> PRT
<213> Escherichia coli

<400> 393
Met Lys Gly Arg Leu Leu Asp Ala Val Pro Leu Ser Ser Leu Thr Gly
1 5 10 15
Val Gly Ala Ala Leu Ser Asn Lys Leu Ala Lys Ile Asn Leu His Thr
20 25 30
Val Gln Asp Leu Leu Leu His Leu Pro Leu Arg Tyr Glu Asp Arg Thr
35 40 45
His Leu Tyr Pro Ile Gly Glu Leu Leu Pro Gly Val Tyr Ala Thr Val
50 55 60
Glu Gly Glu Val Leu Asn Cys Asn Ile Ser Phe Gly Gly Arg Arg Met
65 70 75 80
Met Thr Cys Gln Ile Ser Asp Gly Ser Gly Ile Leu Thr Met Arg Phe
85 90 95
Phe Asn Phe Ser Ala Ala Met Lys Asn Ser Leu Ala Ala Gly Arg Arg
100 105 110
Val Leu Ala Tyr Gly Glu Ala Lys Arg Gly Lys Tyr Gly Ala Glu Met
115 120 125
Ile His Pro Glu Tyr Arg Val Gln Gly Asp Leu Ser Thr Pro Glu Leu
130 135 140
Gln Glu Thr Leu Thr Pro Val Tyr Pro Thr Thr Glu Gly Val Lys Gln
145 150 155 160
Ala Thr Leu Arg Lys Leu Thr Asp Gln Ala Leu Asp Leu Leu Asp Thr
165 170 175
Cys Ala Ile Glu Glu Leu Leu Pro Pro Glu Leu Ser Gln Gly Met Met
180 185 190
Thr Leu Pro Glu Ala Leu Arg Thr Leu His Arg Pro Pro Pro Thr Leu
195 200 205
Gln Leu Ser Asp Leu Glu Thr Gly Gln His Pro Ala Gln Arg Arg Leu
210 215 220
Ile Leu Glu Glu Leu Leu Ala His Asn Leu Ser Met Leu Ala Leu Arg
225 230 235 240
Ala Gly Ala Gln Arg Phe His Ala Gln Pro Leu Ser Ala Asn Asp Thr
245 250 255
Leu Lys Asn Lys Leu Leu Ala Ala Leu Pro Phe Lys Pro Thr Gly Ala
260 265 270
Gln Ala Arg Val Val Ala Glu Ile Glu Arg Asp Met Ala Leu Asp Val
275 280 285
Pro Met Met Arg Leu Val Gln Gly Asp Val Gly Ser Gly Lys Thr Leu
290 295 300
Val Ala Ala Leu Ala Ala Leu Arg Ala Ile Ala His Gly Lys Gln Val
305 310 315 320
Ala Leu Met Ala Pro Thr Glu Leu Leu Ala Glu Gln His Ala Asn Asn
325 330 335
Phe Arg Asn Trp Phe Ala Pro Leu Gly Ile Glu Val Gly Trp Leu Ala
340 345 350

Gly Lys Gln Lys Gly Lys Ala Arg Leu Ala Gln Gln Glu Ala Ile Ala
355 360 365
Ser Gly Gln Val Gln Met Ile Val Gly Thr His Ala Ile Phe Gln Glu
370 375 380
Gln Val Gln Phe Asn Gly Leu Ala Leu Val Ile Ile Asp Glu Gln His
385 390 395 400
Arg Phe Gly Val His Gln Arg Leu Ala Leu Trp Glu Lys Gly Gln Gln
405 410 415
Gln Gly Phe His Pro His Gln Leu Ile Met Thr Ala Thr Pro Ile Pro
420 425 430
Arg Thr Leu Ala Met Thr Ala Tyr Ala Asp Leu Asp Thr Ser Val Ile
435 440 445
Asp Glu Leu Pro Pro Gly Arg Thr Pro Val Thr Thr Val Ala Ile Pro
450 455 460
Asp Thr Arg Arg Thr Asp Ile Ile Asp Arg Val His His Ala Cys Ile
465 470 475 480
Thr Glu Gly Arg Gln Ala Tyr Trp Val Cys Thr Leu Ile Glu Glu Ser
485 490 495
Glu Leu Leu Glu Ala Gln Ala Ala Glu Ala Thr Trp Glu Glu Leu Lys
500 505 510
Leu Ala Leu Pro Glu Leu Asn Val Gly Leu Val His Gly Arg Met Lys
515 520 525
Pro Ala Glu Lys Gln Ala Val Met Ala Ser Phe Lys Gln Gly Glu Leu
530 535 540
His Leu Leu Val Ala Thr Thr Val Ile Glu Val Gly Val Asp Val Pro
545 550 555 560
Asn Ala Ser Leu Met Ile Ile Glu Asn Pro Glu Arg Leu Gly Leu Ala
565 570 575
Gln Leu His Gln Leu Arg Gly Arg Val Gly Arg Gly Ala Val Ala Ser
580 585 590
His Cys Val Leu Leu Tyr Lys Thr Pro Leu Ser Lys Thr Ala Gln Ile
595 600 605
Arg Leu Gln Val Leu Arg Asp Ser Asn Asp Gly Phe Val Ile Ala Gln
610 615 620
Lys Asp Leu Glu Ile Arg Gly Pro Gly Glu Leu Leu Gly Thr Arg Gln
625 630 635 640
Thr Gly Asn Ala Glu Phe Lys Val Ala Asp Leu Leu Arg Asp Gln Ala
645 650 655
Met Ile Pro Glu Val Gln Arg Leu Ala Arg His Ile His Glu Arg Tyr
660 665 670
Pro Gln Gln Ala Lys Ala Leu Ile Glu Arg Trp Met Pro Glu Thr Glu
675 680 685
Arg Tyr Ser Asn Ala
690

<210> 394
<211> 428
<212> PRT
<213> Escherichia coli

<400> 394
Met Lys Thr Ser Leu Phe Lys Ser Leu Tyr Phe Gln Val Leu Thr Ala
1 5 10 15
Ile Ala Ile Gly Ile Leu Leu Gly His Phe Tyr Pro Glu Ile Gly Glu
20 25 30
Gln Met Lys Pro Leu Gly Asp Gly Phe Val Lys Leu Ile Lys Met Ile
35 40 45
Ile Ala Pro Val Ile Phe Cys Thr Val Val Thr Gly Ile Ala Gly Met
50 55 60
Glu Ser Met Lys Ala Val Gly Arg Thr Gly Ala Val Ala Leu Leu Tyr

```
65                  70                  75                  80
Phe Glu Ile Val Ser Thr Ile Ala Leu Ile Ile Gly Leu Ile Ile Val
                85                  90                  95
Asn Val Val Gln Pro Gly Ala Gly Met Asn Val Asp Pro Ala Thr Leu
            100                 105                 110
Asp Ala Lys Ala Val Ala Val Tyr Ala Asp Gln Ala Lys Asp Gln Gly
        115                 120                 125
Ile Val Ala Phe Ile Met Asp Val Ile Pro Ala Ser Val Ile Gly Ala
    130                 135                 140
Phe Ala Ser Gly Asn Ile Leu Gln Val Leu Leu Phe Ala Val Leu Phe
145                 150                 155                 160
Gly Phe Ala Leu His Arg Leu Gly Ser Lys Gly Gln Leu Ile Phe Asn
                165                 170                 175
Val Ile Glu Ser Phe Ser Gln Val Ile Phe Gly Ile Ile Asn Met Ile
            180                 185                 190
Met Arg Leu Ala Pro Ile Gly Ala Phe Gly Ala Met Ala Phe Thr Ile
        195                 200                 205
Gly Lys Tyr Gly Val Gly Thr Leu Val Gln Leu Gly Gln Leu Ile Ile
    210                 215                 220
Cys Phe Tyr Ile Thr Cys Ile Leu Phe Val Val Leu Val Leu Gly Ser
225                 230                 235                 240
Ile Ala Lys Ala Thr Gly Phe Ser Ile Phe Lys Phe Ile Arg Tyr Ile
                245                 250                 255
Arg Glu Glu Leu Leu Ile Val Leu Gly Thr Ser Ser Ser Glu Ser Ala
            260                 265                 270
Leu Pro Arg Met Leu Asp Lys Met Glu Lys Leu Gly Cys Arg Lys Ser
        275                 280                 285
Val Val Gly Leu Val Ile Pro Thr Gly Tyr Ser Phe Asn Leu Asp Gly
    290                 295                 300
Thr Ser Ile Tyr Leu Thr Met Ala Ala Val Phe Ile Ala Gln Ala Thr
305                 310                 315                 320
Asn Ser Gln Met Asp Ile Val His Gln Ile Thr Leu Leu Ile Val Leu
                325                 330                 335
Leu Leu Ser Ser Lys Gly Ala Ala Gly Val Thr Gly Ser Gly Phe Ile
            340                 345                 350
Val Leu Ala Ala Thr Leu Ser Ala Val Gly His Leu Pro Val Ala Gly
        355                 360                 365
Leu Ala Leu Ile Leu Gly Ile Asp Arg Phe Met Ser Glu Ala Arg Ala
    370                 375                 380
Leu Thr Asn Leu Val Gly Asn Gly Val Ala Thr Ile Val Val Ala Lys
385                 390                 395                 400
Trp Val Lys Glu Leu Asp His Lys Lys Leu Asp Asp Val Leu Asn Asn
                405                 410                 415
Arg Ala Pro Asp Gly Lys Thr His Glu Leu Ser Ser
            420                 425
```

<210> 395
<211> 396
<212> PRT
<213> Escherichia coli

<400> 395

```
Met Thr Thr Arg Gln His Ser Ser Phe Ala Ile Val Phe Ile Leu Gly
 1               5                  10                  15
Leu Leu Ala Met Leu Met Pro Leu Ser Ile Asp Met Tyr Leu Pro Ala
            20                  25                  30
Leu Pro Val Ile Ser Ala Gln Phe Gly Val Pro Ala Gly Ser Thr Gln
        35                  40                  45
Met Thr Leu Ser Thr Tyr Ile Leu Gly Phe Ala Leu Gly Gln Leu Ile
    50                  55                  60
```

Tyr Gly Pro Met Ala Asp Ser Phe Gly Arg Lys Pro Val Val Leu Gly
65 70 75 80
Gly Thr Leu Val Phe Ala Ala Ala Ala Val Ala Cys Ala Leu Ala Asn
85 90 95
Thr Ile Asp Gln Leu Ile Val Met Arg Phe Phe His Gly Leu Ala Ala
100 105 110
Ala Ala Ala Ser Val Val Ile Asn Ala Leu Met Arg Asp Ile Tyr Pro
115 120 125
Lys Glu Glu Phe Ser Arg Met Met Ser Phe Val Met Leu Val Thr Thr
130 135 140
Ile Ala Pro Leu Met Ala Pro Ile Val Gly Gly Trp Val Leu Val Trp
145 150 155 160
Leu Ser Trp His Tyr Ile Phe Trp Ile Leu Ala Leu Ala Ala Ile Leu
165 170 175
Ala Ser Ala Met Ile Phe Phe Leu Ile Lys Glu Thr Leu Pro Pro Glu
180 185 190
Arg Arg Gln Pro Phe His Ile Arg Thr Thr Ile Gly Asn Phe Ala Ala
195 200 205
Leu Phe Arg His Lys Arg Val Leu Ser Tyr Met Leu Ala Ser Gly Phe
210 215 220
Ser Phe Ala Gly Met Phe Ser Phe Leu Ser Ala Gly Pro Phe Val Tyr
225 230 235 240
Ile Glu Ile Asn His Val Ala Pro Glu Asn Phe Gly Tyr Tyr Phe Ala
245 250 255
Leu Asn Ile Val Phe Leu Phe Val Met Thr Ile Phe Asn Ser Arg Phe
260 265 270
Val Arg Arg Ile Gly Ala Leu Asn Met Phe Arg Ser Gly Leu Trp Ile
275 280 285
Gln Phe Ile Met Ala Ala Trp Met Val Ile Ser Ala Leu Leu Gly Leu
290 295 300
Gly Phe Trp Ser Leu Val Val Gly Val Ala Ala Phe Val Gly Cys Val
305 310 315 320
Ser Met Val Ser Ser Asn Ala Met Ala Val Ile Leu Asp Glu Phe Pro
325 330 335
His Met Ala Gly Thr Ala Ser Ser Leu Ala Gly Thr Phe Arg Phe Gly
340 345 350
Ile Gly Ala Ile Val Gly Ala Leu Leu Ser Leu Ala Thr Phe Asn Ser
355 360 365
Ala Trp Pro Met Ile Trp Ser Ile Ala Phe Cys Ala Thr Ser Ser Ile
370 375 380
Leu Phe Cys Leu Tyr Ala Ser Arg Pro Lys Lys Arg
385 390 395

<210> 396
<211> 231
<212> PRT
<213> Escherichia coli

<400> 396
Met Arg Leu Asp Lys Phe Ile Ala Gln Gln Leu Gly Val Ser Arg Ala
1 5 10 15
Ile Ala Gly Arg Glu Ile Arg Gly Asn Arg Val Thr Val Asp Gly Glu
20 25 30
Ile Val Arg Asn Ala Ala Phe Lys Leu Leu Pro Glu His Asp Val Ala
35 40 45
Tyr Asp Gly Asn Pro Leu Ala Gln Gln His Gly Pro Arg Tyr Phe Met
50 55 60
Leu Asn Lys Pro Gln Gly Tyr Val Cys Ser Thr Asp Asp Pro Asp His
65 70 75 80
Pro Thr Val Leu Tyr Phe Leu Asp Glu Pro Val Ala Trp Lys Leu His

85 90 95
Ala Ala Gly Arg Leu Asp Ile Asp Thr Thr Gly Leu Val Leu Met Thr
100 105 110
Asp Asp Gly Gln Trp Ser His Arg Ile Thr Ser Pro Arg His His Cys
115 120 125
Glu Lys Thr Tyr Leu Val Thr Leu Glu Ser Pro Val Ala Asp Asp Thr
130 135 140
Ala Glu Gln Phe Ala Lys Gly Val Gln Leu His Asn Glu Lys Asp Leu
145 150 155 160
Thr Lys Pro Ala Val Leu Glu Val Ile Thr Pro Thr Gln Val Arg Leu
165 170 175
Thr Ile Ser Glu Gly Arg Tyr His Gln Val Lys Arg Met Phe Ala Ala
180 185 190
Val Gly Asn His Val Val Glu Leu His Arg Glu Arg Ile Gly Gly Ile
195 200 205
Thr Leu Asp Ala Asp Leu Ala Pro Gly Glu Tyr Arg Pro Leu Thr Glu
210 215 220
Glu Glu Ile Ala Ser Val Val
225 230

<210> 397
<211> 442
<212> PRT
<213> Escherichia coli

<400> 397
Met Lys Lys Ile Glu Cys Ala Cys Asn Phe Leu Met Asp Lys Asp Ala
1 5 10 15
Gln Gly Tyr Ile Asp Leu Ser Asp Leu Asp Leu Thr Ser Cys His Phe
20 25 30
Lys Gly Asp Val Ile Ser Lys Val Ser Phe Leu Ser Ser Asn Leu Gln
35 40 45
His Val Thr Phe Glu Cys Lys Glu Ile Gly Asp Cys Asn Phe Thr Thr
50 55 60
Ala Ile Val Asp Asn Val Ile Phe Arg Cys Arg Arg Leu His Asn Val
65 70 75 80
Ile Phe Ile Lys Ala Ser Gly Glu Cys Val Asp Phe Ser Lys Asn Ile
85 90 95
Leu Asp Thr Val Asp Phe Ser Gln Ser Gln Leu Gly His Ser Asn Phe
100 105 110
Arg Glu Cys Gln Ile Arg Asn Ser Asn Phe Asp Asn Cys Tyr Leu Tyr
115 120 125
Ala Ser His Phe Thr Arg Ala Glu Phe Leu Ser Ala Lys Glu Ile Ser
130 135 140
Phe Ile Lys Ser Asn Leu Thr Ala Val Met Phe Asp Tyr Val Arg Met
145 150 155 160
Ser Thr Gly Asn Phe Lys Asp Cys Ile Thr Glu Gln Leu Glu Leu Thr
165 170 175
Ile Asp Tyr Ser Asp Ile Phe Trp Asn Glu Asp Leu Asp Gly Tyr Ile
180 185 190
Asn Asn Ile Ile Lys Met Ile Asp Thr Leu Pro Asp Asn Ala Met Ile
195 200 205
Leu Lys Ser Val Leu Ala Val Lys Leu Val Met Gln Leu Lys Ile Leu
210 215 220
Asn Ile Val Asn Lys Asn Phe Ile Glu Asn Met Lys Lys Ile Phe Ser
225 230 235 240
His Cys Pro Tyr Ile Lys Asp Pro Ile Ile Arg Ser Tyr Ile His Ser
245 250 255
Asp Glu Asp Asn Lys Phe Asp Asp Phe Met Arg Gln His Arg Phe Ser
260 265 270

Glu Val Asn Phe Asp Thr Gln Gln Met Ile Asp Phe Ile Asn Arg Phe
275 280 285
Asn Thr Asn Lys Trp Leu Ile Asp Lys Asn Asn Asn Phe Phe Ile Gln
290 295 300
Leu Ile Asp Gln Ala Leu Arg Ser Thr Asp Asp Met Ile Lys Ala Asn
305 310 315 320
Val Trp His Leu Tyr Lys Glu Trp Ile Arg Ser Asp Asp Val Ser Pro
325 330 335
Ile Phe Ile Glu Thr Glu Asp Asn Leu Arg Thr Phe Asn Thr Asn Glu
340 345 350
Leu Thr Arg Asn Asp Asn Ile Phe Ile Leu Phe Ser Ser Val Asp Asp
355 360 365
Gly Pro Val Met Val Val Ser Ser Gln Arg Leu His Asp Met Leu Asn
370 375 380
Pro Thr Lys Asp Thr Asn Trp Asn Ser Thr Tyr Ile Tyr Lys Ser Arg
385 390 395 400
His Glu Met Leu Pro Val Asn Leu Thr Gln Glu Thr Leu Phe Ser Ser
405 410 415
Lys Ser His Gly Lys Tyr Ala Leu Phe Pro Ile Phe Thr Ala Ser Trp
420 425 430
Arg Ala His Arg Ile Met Asn Lys Gly Val
435 440

<210> 398
<211> 238
<212> PRT
<213> Escherichia coli

<400> 398
Met Gly Arg Lys Trp Ala Asn Ile Val Ala Lys Lys Thr Ala Lys Asp
1 5 10 15
Gly Ala Thr Ser Lys Ile Tyr Ala Lys Phe Gly Val Glu Ile Tyr Ala
20 25 30
Ala Ala Lys Gln Gly Glu Pro Asp Pro Glu Leu Asn Thr Ser Leu Lys
35 40 45
Phe Val Ile Glu Arg Ala Lys Gln Ala Gln Val Pro Lys His Val Ile
50 55 60
Asp Lys Ala Ile Asp Lys Ala Lys Gly Gly Gly Asp Glu Thr Phe Val
65 70 75 80
Gln Gly Arg Tyr Glu Gly Phe Gly Pro Asn Gly Ser Met Ile Ile Ala
85 90 95
Glu Thr Leu Thr Ser Asn Val Asn Arg Thr Ile Ala Asn Val Arg Thr
100 105 110
Ile Phe Asn Lys Lys Gly Gly Asn Ile Gly Ala Ala Gly Ser Val Ser
115 120 125
Tyr Met Phe Asp Asn Thr Gly Val Ile Val Phe Lys Gly Thr Asp Pro
130 135 140
Asp His Ile Phe Glu Ile Leu Leu Glu Ala Glu Val Asp Val Arg Asp
145 150 155 160
Val Thr Glu Glu Glu Gly Asn Ile Val Ile Tyr Thr Glu Pro Thr Asp
165 170 175
Leu His Lys Gly Ile Ala Ala Leu Lys Ala Ala Gly Ile Thr Glu Phe
180 185 190
Ser Thr Thr Glu Leu Glu Met Ile Ala Gln Ser Glu Val Glu Leu Ser
195 200 205
Pro Glu Asp Leu Glu Ile Phe Glu Gly Leu Val Asp Ala Leu Glu Asp
210 215 220
Asp Asp Asp Val Gln Lys Val Tyr His Asn Val Ala Asn Leu
225 230 235

<210> 399
<211> 261
<212> PRT
<213> Escherichia coli

<400> 399
Met Val Leu Met Ser Glu Thr Lys Asn Glu Leu Glu Asp Leu Leu Glu
1   5   10   15
Lys Ala Ala Thr Glu Pro Ala His Arg Pro Ala Phe Phe Arg Thr Leu
            20   25   30
Leu Glu Ser Thr Val Trp Val Pro Gly Thr Ala Ala Gln Gly Glu Ala
        35   40   45
Val Val Glu Asp Ser Ala Leu Asp Leu Gln His Trp Glu Lys Glu Asp
    50   55   60
Gly Thr Ser Val Ile Pro Phe Phe Thr Ser Leu Glu Ala Leu Gln Gln
65   70   75   80
Ala Val Glu Asp Glu Gln Ala Phe Val Val Met Pro Val Arg Thr Leu
                85   90   95
Phe Glu Met Thr Leu Gly Glu Thr Leu Phe Leu Asn Ala Lys Leu Pro
            100   105   110
Thr Gly Lys Glu Phe Met Pro Arg Glu Ile Ser Leu Leu Ile Gly Glu
        115   120   125
Glu Gly Asn Pro Leu Ser Ser Gln Glu Ile Leu Glu Gly Gly Glu Ser
    130   135   140
Leu Ile Leu Ser Glu Val Ala Glu Pro Pro Ala Gln Met Ile Asp Ser
145   150   155   160
Leu Thr Thr Leu Phe Lys Thr Ile Lys Pro Val Lys Arg Ala Phe Ile
                165   170   175
Cys Ser Ile Lys Glu Asn Glu Glu Ala Gln Pro Asn Leu Leu Ile Gly
            180   185   190
Ile Glu Ala Asp Gly Asp Ile Glu Glu Ile Ile Gln Ala Thr Gly Ser
        195   200   205
Val Ala Thr Asp Thr Leu Pro Gly Asp Glu Pro Ile Asp Ile Cys Gln
    210   215   220
Val Lys Lys Gly Glu Lys Gly Ile Ser His Phe Ile Thr Glu His Ile
225   230   235   240
Ala Pro Phe Tyr Glu Arg Arg Trp Gly Gly Phe Leu Arg Asp Phe Lys
                245   250   255
Gln Asn Arg Ile Ile
            260

<210> 400
<211> 421
<212> PRT
<213> Escherichia coli

<400> 400
Met Leu Thr Lys Lys Lys Trp Ala Leu Phe Ser Leu Leu Thr Leu Cys
1   5   10   15
Gly Gly Thr Ile Tyr Lys Leu Pro Ser Leu Lys Asp Ala Phe Tyr Ile
            20   25   30
Pro Met Gln Glu Tyr Phe His Leu Thr Asn Gly Gln Ile Gly Asn Ala
        35   40   45
Met Ser Val Asn Ser Phe Val Thr Thr Val Gly Phe Phe Leu Ser Ile
    50   55   60
Tyr Phe Ala Asp Lys Leu Pro Arg Arg Tyr Thr Met Ser Phe Ser Leu
65   70   75   80
Ile Ala Thr Gly Leu Leu Gly Val Tyr Leu Thr Thr Met Pro Gly Tyr
                85   90   95

Trp Gly Ile Leu Phe Val Trp Ala Leu Phe Gly Val Thr Cys Asp Met
100 105 110
Met Asn Trp Pro Val Leu Leu Lys Ser Val Ser Arg Leu Gly Asn Ser
115 120 125
Glu Gln Gln Gly Arg Leu Phe Gly Phe Phe Glu Thr Gly Arg Gly Ile
130 135 140
Val Asp Thr Val Val Ala Phe Ser Ala Leu Ala Val Phe Thr Trp Phe
145 150 155 160
Gly Ser Gly Leu Leu Gly Phe Lys Ala Gly Ile Trp Phe Tyr Ser Leu
165 170 175
Ile Val Ile Ala Val Gly Ile Ile Ile Phe Phe Val Leu Asn Asp Lys
180 185 190
Glu Glu Ala Pro Ser Val Glu Val Lys Lys Glu Asp Gly Ala Ser Lys
195 200 205
Asn Thr Ser Met Thr Ser Val Leu Lys Asp Lys Thr Ile Trp Leu Ile
210 215 220
Ala Phe Asn Val Phe Phe Val Tyr Ala Val Tyr Cys Gly Leu Thr Phe
225 230 235 240
Phe Ile Pro Phe Leu Lys Asn Ile Tyr Leu Leu Pro Val Ala Leu Val
245 250 255
Gly Ala Tyr Gly Ile Ile Asn Gln Tyr Cys Leu Lys Met Ile Gly Gly
260 265 270
Pro Ile Gly Gly Met Ile Ser Asp Lys Ile Leu Lys Ser Pro Ser Lys
275 280 285
Tyr Leu Cys Tyr Thr Phe Ile Ile Ser Thr Ala Ala Leu Val Leu Leu
290 295 300
Ile Met Leu Pro His Glu Ser Met Pro Val Tyr Leu Gly Met Ala Cys
305 310 315 320
Thr Leu Gly Phe Gly Ala Ile Val Phe Thr Gln Arg Ala Val Phe Phe
325 330 335
Ala Pro Ile Gly Glu Ala Lys Ile Ala Glu Asn Lys Thr Gly Ala Ala
340 345 350
Met Ala Leu Gly Ser Phe Ile Gly Tyr Ala Pro Ala Met Phe Cys Phe
355 360 365
Ser Leu Tyr Gly Tyr Ile Leu Asp Leu Asn Pro Gly Ile Ile Gly Tyr
370 375 380
Lys Ile Val Phe Gly Ile Met Ala Cys Phe Ala Phe Ser Gly Ala Val
385 390 395 400
Val Ser Val Met Leu Val Lys Arg Ile Ser Gln Arg Lys Lys Glu Met
405 410 415
Leu Ala Ala Glu Ala
420

<210> 401
<211> 770
<212> PRT
<213> Escherichia coli

<400> 401
Met Pro Arg Leu Leu Thr Lys Arg Gly Cys Trp Ile Thr Leu Ala Ala
1 5 10 15
Ala Pro Phe Leu Leu Phe Leu Ala Ala Trp Gly Ala Asp Lys Leu Trp
20 25 30
Pro Leu Pro Leu His Glu Val Asn Pro Ala Arg Val Val Val Ala Gln
35 40 45
Asp Gly Thr Pro Leu Trp Arg Phe Ala Asp Ala Asp Gly Ile Trp Arg
50 55 60
Tyr Pro Val Thr Ile Glu Asp Val Ser Pro Arg Tyr Leu Glu Ala Leu
65 70 75 80
Ile Asn Tyr Glu Asp Arg Trp Phe Trp Lys His Pro Gly Val Asn Pro

```
                85                  90                  95
Phe Ser Val Ala Arg Ala Ala Trp Gln Asp Leu Thr Ser Gly Arg Val
            100                 105                 110
Ile Ser Gly Gly Ser Thr Leu Thr Met Gln Val Ala Arg Leu Leu Asp
        115                 120                 125
Pro His Pro Lys Thr Phe Gly Gly Lys Ile Arg Gln Leu Trp Arg Ala
    130                 135                 140
Leu Gln Leu Glu Trp His Leu Ser Lys Arg Glu Ile Leu Thr Leu Tyr
145                 150                 155                 160
Leu Asn Arg Ala Pro Phe Gly Gly Thr Leu Gln Gly Ile Gly Ala Ala
                165                 170                 175
Ser Trp Ala Tyr Leu Gly Lys Ser Pro Ala Asn Leu Ser Tyr Ser Glu
            180                 185                 190
Ala Ala Met Leu Ala Val Leu Pro Gln Ala Pro Ser Arg Leu Arg Pro
        195                 200                 205
Asp Arg Trp Pro Glu Arg Ala Glu Ala Ala Arg Asn Lys Val Leu Glu
    210                 215                 220
Arg Met Ala Val Gln Gly Val Trp Ser Arg Glu Gln Val Lys Glu Ser
225                 230                 235                 240
Arg Glu Glu Pro Ile Trp Leu Ala Pro Arg Gln Met Pro Gln Leu Ala
                245                 250                 255
Pro Leu Phe Ser Arg Met Met Leu Gly Lys Ser Lys Ser Asp Lys Ile
            260                 265                 270
Thr Thr Thr Leu Asp Ala Gly Leu Gln Arg Arg Leu Glu Glu Leu Ala
        275                 280                 285
Gln Asn Trp Lys Gly Arg Leu Pro Pro Arg Ser Ser Leu Ala Met Ile
    290                 295                 300
Val Val Asp His Thr Asp Met Arg Val Arg Gly Trp Val Gly Ser Val
305                 310                 315                 320
Asp Leu Asn Asp Asp Ser Arg Phe Gly His Val Asp Met Val Asn Ser
                325                 330                 335
Ile Arg Ser Pro Gly Ser Val Leu Lys Pro Phe Val Tyr Gly Leu Ala
            340                 345                 350
Leu Asp Glu Gly Leu Ile His Pro Ala Ser Leu Leu Gln Asp Val Pro
        355                 360                 365
Arg Arg Thr Gly Asp Tyr Arg Pro Gly Asn Phe Asp Ser Gly Phe His
    370                 375                 380
Gly Pro Ile Ser Met Ser Glu Ala Leu Val Arg Ser Leu Asn Leu Pro
385                 390                 395                 400
Ala Val Gln Val Leu Glu Ala Tyr Gly Pro Lys Arg Phe Ala Ala Lys
                405                 410                 415
Leu Arg Asn Val Gly Leu Pro Leu Tyr Leu Pro Asn Gly Ala Ala Pro
            420                 425                 430
Asn Leu Ser Leu Ile Leu Gly Gly Ala Gly Ala Lys Leu Glu Asp Met
        435                 440                 445
Ala Ala Ala Tyr Thr Ala Phe Ala Arg His Gly Lys Ala Gly Lys Leu
    450                 455                 460
Arg Leu Gln Pro Asp Asp Pro Leu Leu Glu Arg Pro Leu Met Ser Ser
465                 470                 475                 480
Gly Ala Ala Trp Ile Ile Arg Arg Ile Met Ala Asp Glu Ala Gln Pro
                485                 490                 495
Leu Pro Asp Ser Ala Leu Pro Arg Val Ala Pro Leu Ala Trp Lys Thr
            500                 505                 510
Gly Thr Ser Tyr Gly Tyr Arg Asp Ala Trp Ala Ile Gly Val Asn Ala
        515                 520                 525
Arg Tyr Val Ile Gly Ile Trp Thr Gly Arg Pro Asp Gly Thr Pro Val
    530                 535                 540
Val Gly Gln Phe Gly Phe Ala Ser Ala Val Pro Leu Leu Asn Gln Val
545                 550                 555                 560
Asn Asn Ile Leu Leu Ser Arg Ser Ala Asn Leu Pro Glu Asp Pro Arg
                565                 570                 575
```

Pro Asn Ser Val Thr Arg Gly Val Ile Cys Trp Pro Gly Gly Gln Ser
580 585 590
Leu Pro Glu Gly Asp Gly Asn Cys Arg Arg Arg Leu Ala Thr Trp Leu
595 600 605
Leu Asp Gly Ser Gln Pro Pro Thr Leu Leu Leu Pro Glu Gln Glu Gly
610 615 620
Ile Asn Gly Ile Arg Phe Pro Ile Trp Leu Asp Glu Asn Gly Lys Arg
625 630 635 640
Val Ala Ala Asp Cys Pro Gln Ala Arg Gln Glu Met Ile Asn Val Trp
645 650 655
Pro Leu Pro Leu Glu Pro Trp Leu Pro Ala Ser Glu Arg Arg Ala Val
660 665 670
Arg Leu Pro Pro Ala Ser Thr Ser Cys Pro Pro Tyr Gly His Asp Ala
675 680 685
Gln Leu Pro Leu Gln Leu Thr Gly Val Arg Asp Gly Ala Ile Ile Lys
690 695 700
Arg Leu Pro Gly Ala Ala Glu Ala Thr Leu Pro Leu Gln Ser Ser Gly
705 710 715 720
Gly Ala Gly Glu Arg Trp Trp Phe Leu Asn Gly Glu Pro Leu Thr Glu
725 730 735
Arg Gly Arg Asn Val Thr Leu His Leu Thr Asp Lys Gly Asp Tyr Gln
740 745 750
Leu Leu Val Met Asp Asp Val Gly Gln Ile Ala Thr Val Lys Phe Val
755 760 765
Met Gln
770

<210> 402
<211> 1653
<212> PRT
<213> Escherichia coli

<400> 402
Met Lys Lys Leu Arg Val Ala Ala Cys Met Leu Met Leu Ala Leu Ala
1 5 10 15
Gly Cys Asp Asn Asn Asp Asn Ala Pro Thr Ala Val Lys Lys Asp Ala
20 25 30
Pro Ser Glu Val Thr Lys Ala Ala Ser Ser Glu Asn Ala Ser Ser Ala
35 40 45
Lys Leu Ser Val Pro Glu Arg Gln Lys Leu Ala Gln Gln Ser Ala Gly
50 55 60
Lys Val Leu Thr Leu Leu Asp Leu Ser Glu Val Gln Leu Asp Gly Ala
65 70 75 80
Ala Thr Leu Val Leu Thr Phe Ser Ile Pro Leu Asp Pro Asp Gln Asp
85 90 95
Phe Ser Arg Val Ile His Val Val Asp Lys Lys Ser Gly Lys Val Asp
100 105 110
Gly Ala Trp Glu Leu Ser Asp Asn Leu Lys Glu Leu Arg Leu Arg His
115 120 125
Leu Glu Pro Lys Arg Asp Leu Ile Val Thr Ile Gly Lys Glu Val Lys
130 135 140
Ala Leu Asn Asn Ala Thr Phe Ser Lys Asp Tyr Glu Lys Thr Ile Thr
145 150 155 160
Thr Arg Asp Ile Gln Pro Ser Val Gly Phe Ala Ser Arg Gly Ser Leu
165 170 175
Leu Pro Gly Lys Val Val Glu Gly Leu Pro Val Met Ala Leu Asn Val
180 185 190
Asn Asn Val Asp Val Asn Phe Phe Arg Val Lys Pro Glu Ser Leu Pro
195 200 205
Ala Phe Ile Ser Gln Trp Glu Tyr Arg Asn Ser Leu Ala Asn Trp Gln

```
    210                 215                 220
Ser Asp Lys Leu Leu Gln Met Ala Asp Leu Val Tyr Thr Gly Arg Phe
225                 230                 235                 240
Asp Leu Asn Pro Ala Arg Asn Thr Arg Glu Lys Leu Leu Leu Pro Leu
                245                 250                 255
Gly Asp Ile Lys Pro Leu Gln Gln Ala Gly Val Tyr Leu Ala Val Met
                260                 265                 270
Asn Gln Ala Gly Arg Tyr Asp Tyr Ser Asn Pro Ala Thr Leu Phe Thr
        275                 280                 285
Leu Ser Asp Ile Gly Val Ser Ala His Arg Tyr His Asn Arg Leu Asp
    290                 295                 300
Ile Phe Thr Gln Ser Leu Glu Asn Gly Ala Ala Gln Gln Gly Ile Glu
305                 310                 315                 320
Val Ser Leu Leu Asn Glu Lys Gly Gln Thr Leu Thr Gln Ala Thr Ser
                325                 330                 335
Asp Ala Gln Gly His Val Gln Leu Glu Asn Asp Lys Asn Ala Ala Leu
                340                 345                 350
Leu Leu Ala Arg Lys Asp Gly Gln Thr Thr Leu Leu Asp Leu Lys Leu
        355                 360                 365
Pro Ala Leu Asp Leu Ala Glu Phe Asn Ile Ala Gly Ala Pro Gly Tyr
    370                 375                 380
Ser Lys Gln Phe Phe Met Phe Gly Pro Arg Asp Leu Tyr Arg Pro Gly
385                 390                 395                 400
Glu Thr Val Ile Leu Asn Gly Leu Leu Arg Asp Ala Asp Gly Lys Ala
                405                 410                 415
Leu Pro Asn Gln Pro Ile Lys Leu Asp Val Ile Lys Pro Asp Gly Gln
                420                 425                 430
Val Leu Arg Ser Val Val Ser Gln Pro Glu Asn Gly Leu Tyr His Phe
        435                 440                 445
Thr Trp Pro Leu Asp Ser Asn Ala Ala Thr Gly Met Trp His Ile Arg
    450                 455                 460
Ala Asn Thr Gly Asp Asn Gln Tyr Arg Met Trp Asp Phe His Val Glu
465                 470                 475                 480
Asp Phe Met Pro Glu Arg Met Ala Leu Asn Leu Thr Gly Glu Lys Thr
                485                 490                 495
Pro Leu Thr Pro Lys Asp Glu Val Lys Phe Ser Val Val Gly Tyr Tyr
                500                 505                 510
Leu Tyr Gly Ala Pro Ala Asn Gly Asn Thr Leu Gln Gly Gln Leu Phe
        515                 520                 525
Leu Arg Pro Leu Arg Glu Ala Val Ser Ala Leu Pro Gly Phe Glu Phe
    530                 535                 540
Gly Asp Ile Ala Ala Glu Asn Leu Ser Arg Thr Leu Asp Glu Val Gln
545                 550                 555                 560
Leu Thr Leu Asp Asp Lys Gly Arg Gly Glu Val Ser Thr Glu Ser Gln
                565                 570                 575
Trp Lys Glu Thr His Ser Pro Leu Gln Val Ile Phe Gln Gly Ser Leu
                580                 585                 590
Leu Glu Ser Gly Gly Arg Pro Val Thr Arg Arg Ala Glu Gln Ala Ile
        595                 600                 605
Trp Pro Ala Asp Ala Leu Pro Gly Ile Arg Pro Gln Phe Ala Ser Lys
    610                 615                 620
Ser Val Tyr Asp Tyr Arg Thr Asp Ser Thr Val Lys Gln Pro Ile Val
625                 630                 635                 640
Asp Glu Gly Ser Asn Ala Ala Phe Asp Ile Val Tyr Ser Asp Ala Gln
                645                 650                 655
Gly Val Lys Lys Ala Val Ser Gly Leu Gln Val Arg Leu Ile Arg Glu
                660                 665                 670
Arg Arg Asp Tyr Tyr Trp Asn Trp Ser Glu Asp Glu Gly Trp Gln Ser
        675                 680                 685
Gln Phe Asp Gln Lys Asp Leu Ile Glu Asn Glu Gln Thr Leu Asp Leu
    690                 695                 700
```

Lys Ala Asp Glu Thr Gly Lys Val Ser Phe Pro Val Glu Trp Gly Ala
705 710 715 720
Tyr Arg Leu Glu Val Lys Ala Pro Asn Glu Ala Val Ser Ser Val Arg
725 730 735
Phe Trp Ala Gly Tyr Ser Trp Gln Asp Asn Ser Asp Gly Ser Gly Ala
740 745 750
Val Arg Pro Asp Arg Val Thr Leu Lys Leu Asp Lys Ala Ser Tyr Arg
755 760 765
Pro Gly Asp Thr Ile Lys Leu His Ile Ala Ala Pro Thr Ala Gly Lys
770 775 780
Gly Tyr Ala Met Val Glu Ser Ser Glu Gly Pro Leu Trp Trp Gln Glu
785 790 795 800
Ile Asp Val Arg Ala Gln Gly Leu Asp Leu Thr Ile Pro Val Asp Lys
805 810 815
Thr Trp Asn Arg His Asp Leu Tyr Leu Ser Thr Leu Val Val Arg Pro
820 825 830
Gly Asp Lys Ser Arg Ser Ala Thr Pro Lys Arg Ala Val Gly Val Leu
835 840 845
His Leu Pro Leu Gly Asp Glu Asn Arg Arg Leu Asp Leu Ala Leu Glu
850 855 860
Thr Pro Ala Lys Met Arg Pro Asn Gln Pro Leu Thr Val Lys Ile Lys
865 870 875 880
Ala Ser Thr Lys Asn Gly Glu Lys Pro Lys Gln Val Asn Val Leu Val
885 890 895
Ser Ala Val Asp Ser Gly Val Leu Asn Ile Thr Asp Tyr Val Thr Pro
900 905 910
Asp Pro Trp Gln Ala Phe Phe Gly Gln Lys Arg Tyr Gly Ala Asp Ile
915 920 925
Tyr Asp Ile Tyr Gly Gln Val Ile Glu Gly Gln Gly Arg Leu Ala Ala
930 935 940
Leu Arg Phe Gly Gly Asp Gly Asp Glu Leu Lys Arg Gly Gly Lys Pro
945 950 955 960
Pro Val Asn His Val Asn Ile Val Val Gln Gln Ala Leu Pro Val Thr
965 970 975
Leu Asn Glu Gln Gly Glu Gly Ser Val Thr Leu Pro Ile Gly Asp Phe
980 985 990
Asn Gly Glu Leu Arg Val Met Ala Gln Ala Trp Thr Ala Asp Asp Phe
995 1000 1005
Gly Ser Asn Glu Ser Lys Val Ile Val Ala Ala Pro Val Ile Ala Glu
1010 1015 1020
Leu Asn Met Pro Arg Phe Met Ala Ser Gly Asp Thr Ser Arg Leu Thr
1025 1030 1035 1040
Leu Asp Ile Thr Asn Leu Thr Asp Lys Pro Gln Lys Leu Asn Val Ala
1045 1050 1055
Leu Thr Ala Ser Gly Leu Leu Glu Leu Val Ser Asp Ser Pro Ala Ala
1060 1065 1070
Val Glu Leu Ala Pro Gly Val Arg Thr Thr Leu Phe Ile Pro Val Arg
1075 1080 1085
Ala Leu Pro Gly Tyr Gly Asp Gly Glu Ile Gln Ala Thr Ile Ser Gly
1090 1095 1100
Leu Ala Leu Pro Gly Glu Thr Val Ala Asp Gln His Lys Gln Trp Lys
1105 1110 1115 1120
Ile Gly Val Arg Pro Ala Phe Pro Ala Gln Thr Val Asn Tyr Gly Thr
1125 1130 1135
Ala Leu Gln Pro Gly Glu Thr Trp Ala Ile Pro Ala Asp Gly Leu Gln
1140 1145 1150
Asn Phe Ser Pro Val Thr Leu Glu Gly Gln Leu Leu Leu Ser Gly Lys
1155 1160 1165
Pro Pro Leu Asn Ile Ala Arg Tyr Ile Lys Glu Leu Lys Ala Tyr Pro
1170 1175 1180
Tyr Gly Cys Leu Glu Gln Thr Ala Ser Gly Leu Phe Pro Ser Leu Tyr

1185 1190 1195 1200
Thr Asn Ala Ala Gln Leu Gln Ala Leu Gly Ile Lys Gly Asp Ser Asp
1205 1210 1215
Glu Lys Arg Arg Ala Ser Val Asp Ile Gly Ile Ser Arg Leu Leu Gln
1220 1225 1230
Met Gln Arg Asp Asn Gly Gly Phe Ala Leu Trp Asp Lys Asn Gly Asp
1235 1240 1245
Glu Glu Tyr Trp Leu Thr Ala Tyr Val Met Asp Phe Leu Val Arg Ala
1250 1255 1260
Gly Glu Gln Gly Tyr Ser Val Pro Thr Asp Ala Ile Asn Arg Gly Asn
1265 1270 1275 1280
Glu Arg Leu Leu Arg Tyr Leu Gln Asp Pro Gly Met Met Ser Ile Pro
1285 1290 1295
Tyr Ala Asp Asn Leu Lys Ala Ser Lys Phe Ala Val Gln Ser Tyr Ala
1300 1305 1310
Ala Leu Val Leu Ala Arg Gln Gln Lys Ala Pro Leu Gly Ala Leu Arg
1315 1320 1325
Glu Ile Trp Glu His Arg Ala Asp Ala Ala Ser Gly Leu Pro Leu Leu
1330 1335 1340
Gln Leu Gly Val Ala Leu Lys Thr Met Gly Asp Ala Thr Arg Gly Glu
1345 1350 1355 1360
Glu Ala Ile Ala Leu Ala Leu Lys Thr Pro Arg Asn Ser Asp Glu Arg
1365 1370 1375
Ile Trp Leu Gly Asp Tyr Gly Ser Ser Leu Arg Asp Asn Ala Leu Met
1380 1385 1390
Leu Ser Leu Leu Glu Glu Asn Lys Leu Leu Pro Asp Glu Gln Tyr Thr
1395 1400 1405
Leu Leu Asn Thr Leu Ser Gln Gln Ala Phe Gly Glu Arg Trp Leu Ser
1410 1415 1420
Thr Gln Glu Ser Asn Ala Leu Phe Leu Ala Ala Arg Thr Ile Gln Asp
1425 1430 1435 1440
Leu Pro Gly Lys Trp Gln Ala Gln Thr Ser Phe Ser Ala Glu Gln Leu
1445 1450 1455
Thr Gly Glu Lys Ala Gln Asn Ser Asn Leu Asn Ser Asp Gln Leu Val
1460 1465 1470
Thr Leu Gln Val Ser Asn Ser Gly Asp Gln Pro Leu Trp Leu Arg Met
1475 1480 1485
Asp Ala Ser Gly Tyr Pro Gln Ser Ala Pro Leu Pro Ala Asn Asn Val
1490 1495 1500
Leu Gln Ile Glu Arg His Ile Leu Gly Thr Asp Gly Lys Ser Lys Ser
1505 1510 1515 1520
Leu Asp Ser Leu Arg Ser Gly Asp Leu Val Leu Val Trp Leu Gln Val
1525 1530 1535
Lys Ala Ser Asn Ser Val Pro Asp Ala Leu Val Val Asp Leu Leu Pro
1540 1545 1550
Ala Gly Leu Glu Leu Glu Asn Gln Asn Leu Ala Asn Gly Ser Ala Ser
1555 1560 1565
Leu Glu Gln Ser Gly Gly Glu Val Gln Asn Leu Leu Asn Gln Met Gln
1570 1575 1580
Gln Ala Ser Ile Lys His Ile Glu Phe Arg Asp Asp Arg Phe Val Ala
1585 1590 1595 1600
Ala Val Ala Val Asp Glu Tyr Gln Pro Val Thr Leu Val Tyr Leu Ala
1605 1610 1615
Arg Ala Val Thr Pro Gly Thr Tyr Gln Val Pro Gln Pro Met Val Glu
1620 1625 1630
Ser Met Tyr Val Pro Gln Trp Arg Ala Thr Gly Ala Ala Glu Asp Leu
1635 1640 1645
Leu Ile Val Arg Pro
1650

<210> 403
<211> 264
<212> PRT
<213> Escherichia coli

<400> 403
Met Asp Leu Leu Tyr Arg Val Lys Thr Leu Trp Ala Ala Leu Arg Gly
1   5   10   15
Asn His Tyr Thr Trp Pro Ala Ile Asp Ile Thr Leu Pro Gly Asn Arg
            20   25   30
His Phe His Leu Ile Gly Ser Ile His Met Gly Ser His Asp Met Ala
        35   40   45
Pro Leu Pro Thr Arg Leu Leu Lys Lys Leu Lys Asn Ala Asp Ala Leu
    50   55   60
Ile Val Glu Ala Asp Val Ser Thr Ser Asp Thr Pro Phe Ala Asn Leu
65   70   75   80
Pro Ala Cys Glu Ala Leu Glu Glu Arg Ile Ser Glu Glu Gln Leu Gln
                85   90   95
Asn Leu Gln His Ile Ser Gln Glu Met Gly Ile Ser Pro Ser Leu Phe
            100   105   110
Ser Thr Gln Pro Leu Trp Gln Ile Ala Met Val Leu Gln Ala Thr Gln
        115   120   125
Ala Gln Lys Leu Gly Leu Arg Ala Glu Tyr Gly Ile Asp Tyr Gln Leu
    130   135   140
Leu Gln Ala Ala Lys Gln Gln His Lys Pro Val Ile Glu Leu Glu Gly
145   150   155   160
Ala Glu Asn Gln Ile Ala Met Leu Leu Gln Leu Pro Asp Lys Gly Leu
                165   170   175
Ala Leu Leu Asp Asp Thr Leu Thr His Trp His Thr Asn Ala Arg Leu
            180   185   190
Leu Gln Gln Met Met Ser Trp Trp Leu Asn Ala Pro Pro Gln Asn Asn
        195   200   205
Asp Ile Thr Leu Pro Asn Thr Phe Ser Gln Ser Leu Tyr Asp Val Leu
    210   215   220
Met His Gln Arg Asn Leu Ala Trp Arg Asp Lys Leu Arg Ala Met Pro
225   230   235   240
Pro Gly Arg Tyr Val Val Ala Val Gly Ala Leu His Leu Tyr Gly Glu
                245   250   255
Gly Asn Leu Pro Gln Met Leu Arg
            260

<210> 404
<211> 322
<212> PRT
<213> Escherichia coli

<400> 404
Met Asp Asn Phe Leu Ala Leu Thr Leu Thr Gly Lys Lys Pro Val Ile
1   5   10   15
Thr Glu Arg Glu Ile Asn Gly Val Arg Trp Arg Trp Leu Gly Asp Gly
            20   25   30
Val Leu Glu Leu Thr Pro Leu Thr Pro Pro Gln Gly Ala Leu Val Ile
        35   40   45
Ser Ala Gly Ile His Gly Asn Glu Thr Ala Pro Val Glu Met Leu Asp
    50   55   60
Ala Leu Leu Gly Ala Ile Ser His Gly Glu Ile Pro Leu Arg Trp Arg
65   70   75   80
Leu Leu Val Ile Leu Gly Asn Pro Pro Ala Leu Lys Gln Gly Lys Arg
                85   90   95
Tyr Cys His Ser Asp Met Asn Arg Met Phe Gly Gly Arg Trp Gln Leu

```
            100                 105                 110
Phe Ala Glu Ser Gly Glu Thr Cys Arg Ala Arg Glu Leu Glu Gln Cys
        115                 120                 125
Leu Glu Asp Phe Tyr Asp Gln Gly Lys Glu Ser Val Arg Trp His Leu
    130                 135                 140
Asp Leu His Thr Ala Ile Arg Gly Ser Leu His Pro Gln Phe Gly Val
145                 150                 155                 160
Leu Pro Gln Arg Asp Ile Pro Trp Asp Glu Lys Phe Leu Thr Trp Leu
                165                 170                 175
Gly Ala Ala Gly Leu Glu Ala Leu Val Phe His Gln Glu Pro Gly Gly
            180                 185                 190
Thr Phe Thr His Phe Ser Ala Arg His Phe Gly Ala Leu Ala Cys Thr
        195                 200                 205
Leu Glu Leu Gly Lys Ala Leu Pro Phe Gly Gln Asn Asp Leu Arg Gln
    210                 215                 220
Phe Ala Val Thr Ala Ser Ala Ile Ala Ala Leu Leu Ser Gly Glu Ser
225                 230                 235                 240
Val Gly Ile Val Arg Thr Pro Pro Leu Arg Tyr Arg Val Val Ser Gln
                245                 250                 255
Ile Thr Arg His Ser Pro Ser Phe Glu Met His Met Ala Ser Asp Thr
            260                 265                 270
Leu Asn Phe Met Pro Phe Glu Lys Gly Thr Leu Leu Ala Gln Asp Gly
        275                 280                 285
Glu Glu Arg Phe Thr Val Thr His Asp Val Glu Tyr Val Leu Phe Pro
    290                 295                 300
Asn Pro Leu Val Ala Leu Gly Leu Arg Ala Gly Leu Met Leu Glu Lys
305                 310                 315                 320
Ile Ser
```

<210> 405
<211> 447
<212> PRT
<213> Escherichia coli

<400> 405

```
Met Asn Ala Trp Glu Val Asn Phe Asp Gly Leu Val Gly Leu Thr His
 1               5                  10                  15
His Tyr Ala Gly Leu Ser Phe Gly Asn Glu Ala Ser Thr Arg His Arg
            20                  25                  30
Phe Gln Val Ser Asn Pro Arg Leu Ala Ala Lys Gln Gly Leu Leu Lys
        35                  40                  45
Met Lys Ala Leu Ala Asp Ala Gly Phe Pro Gln Ala Val Ile Pro Pro
    50                  55                  60
His Glu Arg Pro Phe Ile Pro Val Leu Arg Gln Leu Gly Phe Ser Gly
65                  70                  75                  80
Ser Asp Glu Gln Val Leu Glu Lys Val Ala Arg Gln Ala Pro His Trp
                85                  90                  95
Leu Ser Ser Val Ser Ser Ala Ser Pro Met Trp Val Ala Asn Ala Ala
            100                 105                 110
Thr Ile Ala Pro Ser Ala Asp Thr Leu Asp Gly Lys Val His Leu Thr
        115                 120                 125
Val Ala Asn Leu Asn Asn Lys Phe His Arg Ser Leu Glu Ala Pro Val
    130                 135                 140
Thr Glu Ser Leu Leu Lys Ala Ile Phe Asn Asp Glu Glu Lys Phe Ser
145                 150                 155                 160
Val His Ser Ala Leu Pro Gln Val Ala Leu Leu Gly Asp Glu Gly Ala
                165                 170                 175
Ala Asn His Asn Arg Leu Gly Gly His Tyr Gly Glu Pro Gly Met Gln
            180                 185                 190
```

Leu Phe Val Tyr Gly Arg Glu Glu Gly Asn Asp Thr Arg Pro Ser Arg
195 200 205
Tyr Pro Ala Arg Gln Thr Arg Glu Ala Ser Glu Ala Val Ala Arg Leu
210 215 220
Asn Gln Val Asn Pro Gln Gln Val Ile Phe Ala Gln Gln Asn Pro Asp
225 230 235 240
Val Ile Asp Gln Gly Val Phe His Asn Asp Val Ile Ala Val Ser Asn
245 250 255
Arg Gln Val Leu Phe Cys His Gln Gln Ala Phe Ala Arg Gln Ser Gln
260 265 270
Leu Leu Ala Asn Leu Arg Ala Arg Val Asn Gly Phe Met Ala Ile Glu
275 280 285
Val Pro Ala Thr Gln Val Ser Val Ser Asp Thr Val Ser Thr Tyr Leu
290 295 300
Phe Asn Ser Gln Leu Leu Ser Arg Asp Asp Gly Ser Met Met Leu Val
305 310 315 320
Leu Pro Gln Glu Cys Arg Glu His Ala Gly Val Trp Gly Tyr Leu Asn
325 330 335
Glu Leu Leu Ala Ala Asp Asn Pro Ile Ser Glu Leu Lys Val Phe Asp
340 345 350
Leu Arg Glu Ser Met Ala Asn Gly Gly Gly Pro Ala Cys Leu Arg Leu
355 360 365
Arg Val Val Leu Thr Glu Glu Glu Arg Arg Ala Val Asn Pro Ala Val
370 375 380
Met Met Asn Asp Thr Leu Phe Asn Ala Leu Asn Asp Trp Val Asp Arg
385 390 395 400
Tyr Tyr Arg Asp Arg Leu Thr Ala Ala Asp Leu Ala Asp Pro Gln Leu
405 410 415
Leu Arg Glu Gly Arg Glu Ala Leu Asp Val Leu Ser Gln Leu Leu Asn
420 425 430
Leu Gly Ser Val Tyr Pro Phe Gln Arg Glu Gly Gly Gly Asn Gly
435 440 445

<210> 406
<211> 492
<212> PRT
<213> Escherichia coli

<400> 406
Met Thr Leu Trp Ile Asn Gly Asp Trp Ile Thr Gly Gln Gly Ala Ser
1 5 10 15
Arg Val Lys Arg Asn Pro Val Ser Gly Glu Val Leu Trp Gln Gly Asn
20 25 30
Asp Ala Asp Ala Ala Gln Val Glu Gln Ala Cys Arg Ala Ala Arg Ala
35 40 45
Ala Phe Pro Arg Trp Ala Arg Leu Ser Phe Ala Glu Arg His Ala Val
50 55 60
Val Glu Arg Phe Ala Ala Leu Leu Glu Ser Asn Lys Ala Glu Leu Thr
65 70 75 80
Ala Ile Ile Ala Arg Glu Thr Gly Lys Pro Arg Trp Glu Ala Ala Thr
85 90 95
Glu Val Thr Ala Met Ile Asn Lys Ile Ala Ile Ser Ile Lys Ala Tyr
100 105 110
His Val Arg Thr Gly Glu Gln Arg Ser Glu Met Pro Asp Gly Ala Ala
115 120 125
Ser Leu Arg His Arg Pro His Gly Val Leu Ala Val Phe Gly Pro Tyr
130 135 140
Asn Phe Pro Gly His Leu Pro Asn Gly His Ile Val Pro Ala Leu Leu
145 150 155 160
Ala Gly Asn Thr Ile Ile Phe Lys Pro Ser Glu Leu Thr Pro Trp Ser

165 170 175
Gly Glu Ala Val Met Arg Leu Trp Gln Gln Ala Gly Leu Pro Pro Gly
180 185 190
Val Leu Asn Leu Val Gln Gly Gly Arg Glu Thr Gly Gln Ala Leu Ser
195 200 205
Ala Leu Glu Asp Leu Asp Gly Leu Leu Phe Thr Gly Ser Ala Asn Thr
210 215 220
Gly Tyr Gln Leu His Arg Gln Leu Ser Gly Gln Pro Glu Lys Ile Leu
225 230 235 240
Ala Leu Glu Met Gly Gly Asn Asn Pro Leu Ile Ile Asp Glu Val Ala
245 250 255
Asp Ile Asp Ala Ala Val His Leu Thr Ile Gln Ser Ala Phe Val Thr
260 265 270
Ala Gly Gln Arg Cys Thr Cys Ala Arg Arg Leu Leu Leu Lys Ser Gly
275 280 285
Ala Gln Gly Asp Ala Phe Leu Ala Arg Leu Val Ala Val Ser Gln Arg
290 295 300
Leu Thr Pro Gly Asn Trp Asp Asp Glu Pro Gln Pro Phe Ile Gly Gly
305 310 315 320
Leu Ile Ser Glu Gln Ala Ala Gln Gln Val Val Thr Ala Trp Gln Gln
325 330 335
Leu Glu Ala Met Gly Gly Arg Pro Leu Leu Ala Pro Arg Leu Leu Gln
340 345 350
Ala Gly Thr Ser Leu Leu Thr Pro Gly Ile Ile Glu Met Thr Gly Val
355 360 365
Ala Gly Val Pro Asp Glu Glu Val Phe Gly Pro Leu Leu Arg Val Trp
370 375 380
Arg Tyr Asp Thr Phe Asp Glu Ala Ile Arg Met Ala Asn Asn Thr Arg
385 390 395 400
Phe Gly Leu Ser Cys Gly Leu Val Ser Pro Glu Arg Glu Lys Phe Asp
405 410 415
Gln Leu Leu Leu Glu Ala Arg Ala Gly Ile Val Asn Trp Asn Lys Pro
420 425 430
Leu Thr Gly Ala Ala Ser Thr Ala Pro Phe Gly Gly Ile Gly Ala Ser
435 440 445
Gly Asn His Arg Pro Ser Ala Trp Tyr Ala Ala Asp Tyr Cys Ala Trp
450 455 460
Pro Met Ala Ser Leu Glu Ser Asp Ser Leu Thr Leu Pro Ala Thr Leu
465 470 475 480
Asn Pro Gly Leu Asp Phe Ser Asp Glu Val Val Arg
485 490

<210> 407
<211> 344
<212> PRT
<213> Escherichia coli

<400> 407
Met Met Val Ile Arg Pro Val Glu Arg Ser Asp Val Ser Ala Leu Met
1 5 10 15
Gln Leu Ala Ser Lys Thr Gly Gly Gly Leu Thr Ser Leu Pro Ala Asn
20 25 30
Glu Ala Thr Leu Ser Ala Arg Ile Glu Arg Ala Ile Lys Thr Trp Gln
35 40 45
Gly Glu Leu Pro Lys Ser Glu Gln Gly Tyr Val Phe Val Leu Glu Asp
50 55 60
Ser Glu Thr Gly Thr Val Ala Gly Ile Cys Ala Ile Glu Val Ala Val
65 70 75 80
Gly Leu Asn Asp Pro Trp Tyr Asn Tyr Arg Val Gly Thr Leu Val His
85 90 95

```
Ala Ser Lys Glu Leu Asn Val Tyr Asn Ala Leu Pro Thr Leu Phe Leu
            100                 105                 110
Ser Asn Asp His Thr Gly Ser Ser Glu Leu Cys Thr Leu Phe Leu Asp
        115                 120                 125
Pro Asp Trp Arg Lys Glu Gly Asn Gly Tyr Leu Leu Ser Lys Ser Arg
    130                 135                 140
Phe Met Phe Met Ala Ala Phe Arg Asp Lys Phe Asn Asp Lys Val Val
145                 150                 155                 160
Ala Glu Met Arg Gly Val Ile Asp Glu His Gly Tyr Ser Pro Phe Trp
                165                 170                 175
Gln Ser Leu Gly Lys Arg Phe Phe Ser Met Asp Phe Ser Arg Ala Asp
            180                 185                 190
Phe Leu Cys Gly Thr Gly Gln Lys Ala Phe Ile Ala Glu Leu Met Pro
        195                 200                 205
Lys His Pro Ile Tyr Thr His Phe Leu Ser Gln Glu Ala Gln Asp Val
    210                 215                 220
Ile Gly Gln Val His Pro Gln Thr Ala Pro Ala Arg Ala Val Leu Glu
225                 230                 235                 240
Lys Glu Gly Phe Arg Tyr Arg Asn Tyr Ile Asp Ile Phe Asp Gly Gly
                245                 250                 255
Pro Thr Leu Glu Cys Asp Ile Asp Arg Val Arg Ala Ile Arg Lys Ser
            260                 265                 270
Arg Leu Val Glu Val Ala Glu Gly Gln Pro Ala Gln Gly Asp Phe Pro
        275                 280                 285
Ala Cys Leu Val Ala Asn Glu Asn Tyr His His Phe Arg Val Val Leu
    290                 295                 300
Val Arg Thr Asp Pro Ala Thr Glu Arg Leu Ile Leu Thr Ala Ala Gln
305                 310                 315                 320
Leu Asp Ala Leu Lys Cys His Ala Gly Asp Arg Val Arg Leu Val Arg
                325                 330                 335
Leu Cys Ala Glu Glu Lys Thr Ala
            340


&lt;210&gt; 408
&lt;211&gt; 406
&lt;212&gt; PRT
&lt;213&gt; Escherichia coli

&lt;400&gt; 408
Met Ser Gln Pro Ile Thr Arg Glu Asn Phe Asp Glu Trp Met Ile Pro
 1               5                  10                  15
Val Tyr Ala Pro Ala Pro Phe Ile Pro Val Arg Gly Glu Gly Ser Arg
            20                  25                  30
Leu Trp Asp Gln Gln Gly Lys Glu Tyr Ile Asp Phe Ala Gly Gly Ile
        35                  40                  45
Ala Val Asn Ala Leu Gly His Ala His Pro Glu Leu Arg Glu Ala Leu
    50                  55                  60
Asn Glu Gln Ala Ser Lys Phe Trp His Thr Gly Asn Gly Tyr Thr Asn
65                  70                  75                  80
Glu Pro Val Leu Arg Leu Ala Lys Lys Leu Ile Asp Ala Thr Phe Ala
                85                  90                  95
Asp Arg Val Phe Phe Cys Asn Ser Gly Ala Glu Ala Asn Glu Ala Ala
            100                 105                 110
Leu Lys Leu Ala Arg Lys Phe Ala His Asp Arg Tyr Gly Ser His Lys
        115                 120                 125
Ser Gly Ile Val Ala Phe Lys Asn Ala Phe His Gly Arg Thr Leu Phe
    130                 135                 140
Thr Val Ser Ala Gly Gly Gln Pro Ala Tyr Ser Gln Asp Phe Ala Pro
145                 150                 155                 160
Leu Pro Ala Asp Ile Arg His Ala Ala Tyr Asn Asp Ile Asn Ser Ala
```

```
                165                 170                 175
Ser Ala Leu Ile Asp Asp Ser Thr Cys Ala Val Ile Val Glu Pro Ile
            180                 185                 190
Gln Gly Glu Gly Gly Val Val Pro Ala Ser Asn Ala Phe Leu Gln Gly
        195                 200                 205
Leu Arg Glu Leu Cys Asn Arg His Asn Ala Leu Leu Ile Phe Asp Glu
    210                 215                 220
Val Gln Thr Gly Val Gly Arg Thr Gly Glu Leu Tyr Ala Tyr Met His
225                 230                 235                 240
Tyr Gly Val Thr Pro Asp Leu Leu Thr Thr Ala Lys Ala Leu Gly Gly
                245                 250                 255
Gly Phe Pro Val Gly Ala Leu Leu Ala Thr Glu Glu Cys Ala Arg Val
            260                 265                 270
Met Thr Val Gly Thr His Gly Thr Thr Tyr Gly Gly Asn Pro Leu Ala
        275                 280                 285
Ser Ala Val Ala Gly Lys Val Leu Glu Leu Ile Asn Thr Pro Glu Met
    290                 295                 300
Leu Asn Gly Val Lys Gln Arg His Asp Trp Phe Val Glu Arg Leu Asn
305                 310                 315                 320
Thr Ile Asn His Arg Tyr Gly Leu Phe Ser Glu Val Arg Gly Leu Gly
                325                 330                 335
Leu Leu Ile Gly Cys Val Leu Asn Ala Asp Tyr Ala Gly Gln Ala Lys
            340                 345                 350
Gln Ile Ser Gln Glu Ala Ala Lys Ala Gly Val Met Val Leu Ile Ala
        355                 360                 365
Gly Gly Asn Val Val Arg Phe Ala Pro Ala Leu Asn Val Ser Glu Glu
    370                 375                 380
Glu Val Thr Thr Gly Leu Asp Arg Phe Ala Ala Ala Cys Glu His Phe
385                 390                 395                 400
Val Ser Arg Gly Ser Ser
                405
```

<210> 409
<211> 1048
<212> PRT
<213> Escherichia coli

<400> 409

```
Met Lys Ile Leu Ser Leu Arg Leu Lys Asn Leu Asn Ser Leu Lys Gly
 1               5                  10                  15
Glu Trp Lys Ile Asp Phe Thr Arg Glu Pro Phe Ala Ser Asn Gly Leu
            20                  25                  30
Phe Ala Ile Thr Gly Pro Thr Gly Ala Gly Lys Thr Thr Leu Leu Asp
        35                  40                  45
Ala Ile Cys Leu Ala Leu Tyr His Glu Thr Pro Arg Leu Ser Asn Val
    50                  55                  60
Ser Gln Ser Gln Asn Asp Leu Met Thr Arg Asp Thr Ala Glu Cys Leu
65                  70                  75                  80
Ala Glu Val Glu Phe Glu Val Lys Gly Glu Ala Tyr Arg Ala Phe Trp
                85                  90                  95
Ser Gln Asn Arg Ala Arg Asn Gln Pro Asp Gly Asn Leu Gln Val Pro
            100                 105                 110
Arg Val Glu Leu Ala Arg Cys Ala Asp Gly Lys Ile Leu Ala Asp Lys
        115                 120                 125
Val Lys Asp Lys Leu Glu Leu Thr Ala Thr Leu Thr Gly Leu Asp Tyr
    130                 135                 140
Gly Arg Phe Thr Arg Ser Met Leu Leu Ser Gln Gly Gln Phe Ala Ala
145                 150                 155                 160
Phe Leu Asn Ala Lys Pro Lys Glu Arg Ala Glu Leu Leu Glu Glu Leu
                165                 170                 175
```

Thr Gly Thr Glu Ile Tyr Gly Gln Ile Ser Ala Met Val Phe Glu Gln
180 185 190
His Lys Ser Ala Arg Thr Glu Leu Glu Lys Leu Gln Ala Gln Ala Ser
195 200 205
Gly Val Thr Leu Leu Thr Pro Glu Gln Val Gln Ser Leu Thr Ala Ser
210 215 220
Leu Gln Val Leu Thr Asp Glu Glu Lys Gln Leu Ile Thr Ala Gln Gln
225 230 235 240
Gln Glu Gln Gln Ser Leu Asn Trp Leu Thr Arg Gln Asp Glu Leu Gln
245 250 255
Gln Glu Ala Ser Arg Arg Gln Gln Ala Leu Gln Gln Ala Leu Ala Glu
260 265 270
Glu Glu Lys Ala Gln Pro Gln Leu Ala Ala Leu Ser Leu Ala Gln Pro
275 280 285
Ala Arg Asn Leu Arg Pro His Trp Glu Arg Ile Ala Glu His Ser Ala
290 295 300
Ala Leu Ala His Ile Arg Gln Gln Ile Glu Glu Val Asn Thr Arg Leu
305 310 315 320
Gln Ser Thr Met Ala Leu Arg Ala Ser Ile Arg His His Ala Ala Lys
325 330 335
Gln Ser Ala Glu Leu Gln Gln Gln Gln Gln Ser Leu Asn Thr Trp Leu
340 345 350
Gln Glu His Asp Arg Phe Arg Gln Trp Asn Asn Glu Pro Ala Gly Trp
355 360 365
Arg Ala Gln Phe Ser Gln Gln Thr Ser Asp Arg Glu His Leu Arg Gln
370 375 380
Trp Gln Gln Gln Leu Thr His Ala Glu Gln Lys Leu Asn Ala Leu Ala
385 390 395 400
Ala Ile Thr Leu Thr Leu Thr Ala Asp Glu Val Ala Thr Ala Leu Ala
405 410 415
Gln His Ala Glu Gln Arg Pro Leu Arg Gln His Leu Val Ala Leu His
420 425 430
Gly Gln Ile Val Pro Gln Gln Lys Arg Leu Ala Gln Leu Gln Val Ala
435 440 445
Ile Gln Asn Val Thr Gln Glu Gln Thr Gln Arg Asn Ala Ala Leu Asn
450 455 460
Glu Met Arg Gln Arg Tyr Lys Glu Lys Thr Gln Gln Leu Ala Asp Val
465 470 475 480
Lys Thr Ile Cys Glu Gln Glu Ala Arg Ile Lys Thr Leu Glu Ala Gln
485 490 495
Arg Ala Gln Leu Gln Ala Gly Gln Pro Cys Pro Leu Cys Gly Ser Thr
500 505 510
Ser His Pro Ala Val Glu Ala Tyr Gln Ala Leu Glu Pro Gly Val Asn
515 520 525
Gln Ser Arg Leu Leu Ala Leu Glu Asn Glu Val Lys Lys Leu Gly Glu
530 535 540
Glu Gly Ala Thr Leu Arg Gly Gln Leu Asp Ala Ile Thr Lys Gln Leu
545 550 555 560
Gln Arg Asp Glu Asn Glu Ala Gln Ser Leu Arg Gln Asp Glu Gln Ala
565 570 575
Leu Thr Gln Gln Trp Gln Ala Val Thr Ala Ser Leu Asn Ile Thr Leu
580 585 590
Gln Pro Leu Asp Asp Ile Gln Pro Trp Leu Asp Ala Gln Asp Glu His
595 600 605
Glu Arg Gln Leu Arg Leu Leu Ser Gln Arg His Glu Leu Gln Gly Gln
610 615 620
Ile Ala Ala His Asn Gln Gln Ile Ile Gln Tyr Gln Gln Gln Ile Glu
625 630 635 640
Gln Arg Gln Gln Leu Leu Leu Thr Thr Leu Thr Gly Tyr Ala Leu Thr
645 650 655
Leu Pro Gln Glu Asp Glu Glu Glu Ser Trp Leu Ala Thr Arg Gln Gln

```
            660                 665                 670
Glu Ala Gln Ser Trp Gln Gln Arg Gln Asn Glu Leu Thr Ala Leu Gln
        675                 680                 685
Asn Arg Ile Gln Gln Leu Thr Pro Ile Leu Glu Thr Leu Pro Gln Ser
    690                 695                 700
Asp Glu Leu Pro His Cys Glu Glu Thr Val Val Leu Glu Asn Trp Arg
705                 710                 715                 720
Gln Val His Glu Gln Cys Leu Ala Leu His Ser Gln Gln Gln Thr Leu
                725                 730                 735
Gln Gln Gln Asp Val Leu Ala Ala Gln Ser Leu Gln Lys Ala Gln Ala
            740                 745                 750
Gln Phe Asp Thr Ala Leu Gln Ala Ser Val Phe Asp Asp Gln Gln Ala
        755                 760                 765
Phe Leu Ala Ala Leu Met Asp Glu Gln Thr Leu Thr Gln Leu Glu Gln
    770                 775                 780
Leu Lys Gln Asn Leu Glu Asn Gln Arg Arg Gln Ala Gln Thr Leu Val
785                 790                 795                 800
Thr Gln Thr Ala Glu Thr Leu Ala Gln His Gln Gln His Arg Pro Asp
                805                 810                 815
Asp Gly Leu Ala Leu Thr Val Thr Val Glu Gln Ile Gln Gln Glu Leu
            820                 825                 830
Ala Gln Thr His Gln Lys Leu Arg Glu Asn Thr Thr Ser Gln Gly Glu
        835                 840                 845
Ile Arg Gln Gln Leu Lys Gln Asp Ala Asp Asn Arg Gln Gln Gln Gln
    850                 855                 860
Thr Leu Met Gln Gln Ile Ala Gln Met Thr Gln Gln Val Glu Asp Trp
865                 870                 875                 880
Gly Tyr Leu Asn Ser Leu Ile Gly Ser Lys Glu Gly Asp Lys Phe Arg
                885                 890                 895
Lys Phe Ala Gln Gly Leu Thr Leu Asp Asn Leu Val His Leu Ala Asn
            900                 905                 910
Gln Gln Leu Thr Arg Leu His Gly Arg Tyr Leu Leu Gln Arg Lys Ala
        915                 920                 925
Ser Glu Ala Leu Glu Val Glu Val Val Asp Thr Trp Gln Ala Asp Ala
    930                 935                 940
Val Arg Asp Thr Arg Thr Leu Ser Gly Gly Glu Ser Phe Leu Val Ser
945                 950                 955                 960
Leu Ala Leu Ala Leu Ala Leu Ser Asp Leu Val Ser His Lys Thr Arg
                965                 970                 975
Ile Asp Ser Leu Phe Leu Asp Glu Gly Phe Gly Thr Leu Asp Ser Glu
            980                 985                 990
Thr Leu Asp Thr Ala Leu Asp Ala Leu Asp Ala Leu Asn Ala Ser Gly
        995                 1000                1005
Lys Thr Ile Gly Val Ile Ser His Val Glu Ala Met Lys Glu Arg Ile
    1010                1015                1020
Pro Val Gln Ile Lys Val Lys Lys Ile Asn Gly Leu Gly Tyr Ser Lys
1025                1030                1035                1040
Leu Glu Ser Thr Phe Ala Val Lys
                1045
```

<210> 410
<211> 400
<212> PRT
<213> Escherichia coli

<400> 410

```
Met Arg Ile Leu His Thr Ser Asp Trp His Leu Gly Gln Asn Phe Tyr
 1               5                  10                  15
Ser Lys Ser Arg Glu Ala Glu His Gln Ala Phe Leu Asp Trp Leu Leu
            20                  25                  30
```

Glu Thr Ala Gln Thr His Gln Val Asp Ala Ile Ile Val Ala Gly Asp
35 40 45
Val Phe Asp Thr Gly Ser Pro Pro Ser Tyr Ala Arg Thr Leu Tyr Asn
50 55 60
Arg Phe Val Val Asn Leu Gln Gln Thr Gly Cys His Leu Val Val Leu
65 70 75 80
Ala Gly Asn His Asp Ser Val Ala Thr Leu Asn Glu Ser Arg Asp Ile
85 90 95
Met Ala Phe Leu Asn Thr Thr Val Val Ala Ser Ala Gly His Ala Pro
100 105 110
Gln Ile Leu Pro Arg Arg Asp Gly Thr Pro Gly Ala Val Leu Cys Pro
115 120 125
Ile Pro Phe Leu Arg Pro Arg Asp Ile Ile Thr Ser Gln Ala Gly Leu
130 135 140
Asn Gly Ile Glu Lys Gln Gln His Leu Leu Ala Ala Ile Thr Asp Tyr
145 150 155 160
Tyr Gln Gln His Tyr Ala Asp Ala Cys Lys Leu Arg Gly Asp Gln Pro
165 170 175
Leu Pro Ile Ile Ala Thr Gly His Leu Thr Thr Val Gly Ala Ser Lys
180 185 190
Ser Asp Ala Val Arg Asp Ile Tyr Ile Gly Thr Leu Asp Ala Phe Pro
195 200 205
Ala Gln Asn Phe Pro Pro Ala Asp Tyr Ile Ala Leu Gly His Ile His
210 215 220
Arg Ala Gln Ile Ile Gly Gly Met Glu His Val Arg Tyr Cys Gly Ser
225 230 235 240
Pro Ile Pro Leu Ser Phe Asp Glu Cys Gly Lys Ser Lys Tyr Val His
245 250 255
Leu Val Thr Phe Ser Asn Gly Lys Leu Glu Ser Val Glu Asn Leu Asn
260 265 270
Val Pro Val Thr Gln Pro Met Ala Val Leu Lys Gly Asp Leu Ala Ser
275 280 285
Ile Thr Ala Gln Leu Glu Gln Trp Arg Asp Val Ser Gln Glu Pro Pro
290 295 300
Val Trp Leu Asp Ile Glu Ile Thr Thr Asp Glu Tyr Leu His Asp Ile
305 310 315 320
Gln Arg Lys Ile Gln Ala Leu Thr Glu Ser Leu Pro Val Glu Val Leu
325 330 335
Leu Val Arg Arg Ser Arg Glu Gln Arg Glu Arg Val Leu Ala Ser Gln
340 345 350
Gln Arg Glu Thr Leu Ser Glu Leu Ser Val Glu Glu Val Phe Asn Arg
355 360 365
Arg Leu Ala Leu Glu Glu Leu Asp Glu Ser Gln Gln Gln Arg Leu Gln
370 375 380
His Leu Phe Thr Thr Thr Leu His Thr Leu Ala Gly Glu His Glu Ala
385 390 395 400

<210> 411
<211> 254
<212> PRT
<213> Escherichia coli

<400> 411
Met Asn Asp Leu Phe Ser Leu Ala Gly Lys Asn Ile Leu Ile Thr Gly
1 5 10 15
Ser Ala Gln Gly Ile Gly Phe Leu Leu Ala Thr Gly Leu Gly Lys Tyr
20 25 30
Gly Ala Gln Ile Ile Ile Asn Asp Ile Thr Ala Glu Arg Ala Glu Leu
35 40 45
Ala Val Glu Lys Leu His Gln Glu Gly Ile Gln Ala Val Ala Ala Pro

```
    50                  55                  60
Phe Asn Val Thr His Lys His Glu Ile Asp Ala Ala Val Glu His Ile
65                  70                  75                  80
Glu Lys Asp Ile Gly Pro Ile Asp Val Leu Val Asn Asn Ala Gly Ile
                85                  90                  95
Gln Arg Arg His Pro Phe Thr Glu Phe Pro Glu Gln Glu Trp Asn Asp
            100                 105                 110
Val Ile Ala Val Asn Gln Thr Ala Val Phe Leu Val Ser Gln Ala Val
        115                 120                 125
Thr Arg His Met Val Glu Arg Lys Ala Gly Lys Val Ile Asn Ile Cys
    130                 135                 140
Ser Met Gln Ser Glu Leu Gly Arg Asp Thr Ile Thr Pro Tyr Ala Ala
145                 150                 155                 160
Ser Lys Gly Ala Val Lys Met Leu Thr Arg Gly Met Cys Val Glu Leu
                165                 170                 175
Ala Arg His Asn Ile Gln Val Asn Gly Ile Ala Pro Gly Tyr Phe Lys
            180                 185                 190
Thr Glu Met Thr Lys Ala Leu Val Glu Asp Glu Ala Phe Thr Ala Trp
        195                 200                 205
Leu Cys Lys Arg Thr Pro Ala Ala Arg Trp Gly Asp Pro Gln Glu Leu
    210                 215                 220
Ile Gly Ala Ala Val Phe Leu Ser Ser Lys Ala Ser Asp Phe Val Asn
225                 230                 235                 240
Gly His Leu Leu Phe Val Asp Gly Gly Met Leu Val Ala Val
                245                 250


<210> 412
<211> 343
<212> PRT
<213> Escherichia coli

<400> 412
Met Gln Val Lys Thr Gln Ser Cys Val Val Ala Gly Lys Lys Thr Val
 1               5                  10                  15
Ala Val Thr Glu Gln Thr Ile Asp Trp Asn Asn Asn Gly Thr Leu Val
            20                  25                  30
Gln Ile Thr Arg Gly Gly Ile Cys Gly Ser Asp Leu His Tyr Tyr Gln
        35                  40                  45
Glu Gly Lys Val Gly Asn Phe Met Ile Lys Ala Pro Met Val Leu Gly
    50                  55                  60
His Glu Val Ile Gly Lys Val Ile His Ser Asp Ser Ser Glu Leu His
65                  70                  75                  80
Glu Gly Gln Thr Val Ala Ile Asn Pro Ser Lys Pro Cys Gly His Cys
                85                  90                  95
Lys Tyr Cys Ile Glu His Asn Glu Asn Gln Cys Thr Asp Met Arg Phe
            100                 105                 110
Phe Gly Ser Ala Met Tyr Phe Pro His Val Asp Gly Gly Phe Thr Arg
        115                 120                 125
Tyr Lys Met Val Glu Thr Ser Gln Cys Val Pro Tyr Pro Ala Lys Ala
    130                 135                 140
Asp Glu Lys Val Met Ala Phe Ala Glu Pro Leu Ala Val Ala Ile His
145                 150                 155                 160
Ala Ala His Gln Ala Gly Glu Leu Gln Gly Lys Arg Val Phe Ile Ser
                165                 170                 175
Gly Val Gly Pro Ile Gly Cys Leu Ile Val Ser Ala Val Lys Thr Leu
            180                 185                 190
Gly Ala Ala Glu Ile Val Cys Ala Asp Val Ser Pro Arg Ser Leu Ser
        195                 200                 205
Leu Gly Lys Glu Met Gly Ala Asp Val Leu Val Asn Pro Gln Asn Asp
    210                 215                 220
```

Asp Met Asp His Trp Lys Ala Glu Lys Gly Tyr Phe Asp Val Ser Phe
225 230 235 240
Glu Val Ser Gly His Pro Ser Ser Val Asn Thr Cys Leu Glu Val Thr
245 250 255
Arg Ala Arg Gly Val Met Val Gln Val Gly Met Gly Gly Ala Met Ala
260 265 270
Glu Phe Pro Met Met Thr Leu Ile Gly Lys Glu Ile Ser Leu Arg Gly
275 280 285
Ser Phe Arg Phe Thr Ser Glu Phe Asn Thr Ala Val Ser Trp Leu Ala
290 295 300
Asn Gly Val Ile Asn Pro Leu Pro Leu Leu Ser Ala Glu Tyr Pro Phe
305 310 315 320
Thr Asp Leu Glu Glu Ala Leu Arg Phe Ala Gly Asp Lys Thr Gln Ala
325 330 335
Ala Lys Val Gln Leu Val Phe
340

<210> 413
<211> 548
<212> PRT
<213> Escherichia coli

<400> 413
Met Asp Ser Gln Arg Asn Leu Leu Val Ile Ala Leu Leu Phe Val Ser
1 5 10 15
Phe Met Ile Trp Gln Ala Trp Glu Gln Asp Lys Asn Pro Gln Pro Gln
20 25 30
Ala Gln Gln Thr Thr Gln Thr Thr Thr Thr Ala Ala Gly Ser Ala Ala
35 40 45
Asp Gln Gly Val Pro Ala Ser Gly Gln Gly Lys Leu Ile Ser Val Lys
50 55 60
Thr Asp Val Leu Asp Leu Thr Ile Asn Thr Arg Gly Gly Asp Val Glu
65 70 75 80
Gln Ala Leu Leu Pro Ala Tyr Pro Lys Glu Leu Asn Ser Thr Gln Pro
85 90 95
Phe Gln Leu Leu Glu Thr Ser Pro Gln Phe Ile Tyr Gln Ala Gln Ser
100 105 110
Gly Leu Thr Gly Arg Asp Gly Pro Asp Asn Pro Ala Asn Gly Pro Arg
115 120 125
Pro Leu Tyr Asn Val Glu Lys Asp Ala Tyr Val Leu Ala Glu Gly Gln
130 135 140
Asn Glu Leu Gln Val Pro Met Thr Tyr Thr Asp Ala Ala Gly Asn Thr
145 150 155 160
Phe Thr Lys Thr Phe Val Leu Lys Arg Gly Asp Tyr Ala Val Asn Val
165 170 175
Asn Tyr Asn Val Gln Asn Ala Gly Glu Lys Pro Leu Glu Ile Ser Ser
180 185 190
Phe Gly Gln Leu Lys Gln Ser Ile Thr Leu Pro Pro His Leu Asp Thr
195 200 205
Gly Ser Ser Asn Phe Ala Leu His Thr Phe Arg Gly Ala Ala Tyr Ser
210 215 220
Thr Pro Asp Glu Lys Tyr Glu Lys Tyr Lys Phe Asp Thr Ile Ala Asp
225 230 235 240
Asn Glu Asn Leu Asn Ile Ser Ser Lys Gly Gly Trp Val Ala Met Leu
245 250 255
Gln Gln Tyr Phe Ala Thr Ala Trp Ile Pro His Asn Asp Gly Thr Asn
260 265 270
Asn Phe Tyr Thr Ala Asn Leu Gly Asn Gly Ile Ala Ala Ile Gly Tyr
275 280 285
Lys Ser Gln Pro Val Leu Val Gln Pro Gly Gln Thr Gly Ala Met Asn

290 295 300
Ser Thr Leu Trp Val Gly Pro Glu Ile Gln Asp Lys Met Ala Ala Val
305 310 315 320
Ala Pro His Leu Asp Leu Thr Val Asp Tyr Gly Trp Leu Trp Phe Ile
325 330 335
Ser Gln Pro Leu Phe Lys Leu Leu Lys Trp Ile His Ser Phe Val Gly
340 345 350
Asn Trp Gly Phe Ser Ile Ile Ile Ile Thr Phe Ile Val Arg Gly Ile
355 360 365
Met Tyr Pro Leu Thr Lys Ala Gln Tyr Thr Ser Met Ala Lys Met Arg
370 375 380
Met Leu Gln Pro Lys Ile Gln Ala Met Arg Glu Arg Leu Gly Asp Asp
385 390 395 400
Lys Gln Arg Ile Ser Gln Glu Met Met Ala Leu Tyr Lys Ala Glu Lys
405 410 415
Val Asn Pro Leu Gly Gly Cys Phe Pro Leu Leu Ile Gln Met Pro Ile
420 425 430
Phe Leu Ala Leu Tyr Tyr Met Leu Met Gly Ser Val Glu Leu Arg Gln
435 440 445
Ala Pro Phe Ala Leu Trp Ile His Asp Leu Ser Ala Gln Asp Pro Tyr
450 455 460
Tyr Ile Leu Pro Ile Leu Met Gly Val Thr Met Phe Phe Ile Gln Lys
465 470 475 480
Met Ser Pro Thr Thr Val Thr Asp Pro Met Gln Gln Lys Ile Met Thr
485 490 495
Phe Met Pro Val Ile Phe Thr Val Phe Phe Leu Trp Phe Pro Ser Gly
500 505 510
Leu Val Leu Tyr Tyr Ile Val Ser Asn Leu Val Thr Ile Ile Gln Gln
515 520 525
Gln Leu Ile Tyr Arg Gly Leu Glu Lys Arg Gly Leu His Ser Arg Glu
530 535 540
Lys Lys Lys Ser
545

<210> 414
<211> 542
<212> PRT
<213> Escherichia coli

<400> 414
Met Lys Thr Arg Asp Ser Gln Ser Ser Asp Val Ile Ile Ile Gly Gly
1 5 10 15
Gly Ala Thr Gly Ala Gly Ile Ala Arg Asp Cys Ala Leu Arg Gly Leu
20 25 30
Arg Val Ile Leu Val Glu Arg His Asp Ile Ala Thr Gly Ala Thr Gly
35 40 45
Arg Asn His Gly Leu Leu His Ser Gly Ala Arg Tyr Ala Val Thr Asp
50 55 60
Ala Glu Ser Ala Arg Glu Cys Ile Ser Glu Asn Gln Ile Leu Lys Arg
65 70 75 80
Ile Ala Arg His Cys Val Glu Pro Thr Asn Gly Leu Phe Ile Thr Leu
85 90 95
Pro Glu Asp Asp Leu Ser Phe Gln Ala Thr Phe Ile Arg Ala Cys Glu
100 105 110
Glu Ala Gly Ile Ser Ala Glu Ala Ile Asp Pro Gln Gln Ala Arg Ile
115 120 125
Ile Glu Pro Ala Val Asn Pro Ala Leu Ile Gly Ala Val Lys Val Pro
130 135 140
Asp Gly Thr Val Asp Pro Phe Arg Leu Thr Ala Ala Asn Met Leu Asp
145 150 155 160

Ala Lys Glu His Gly Ala Val Ile Leu Thr Ala His Glu Val Thr Gly
165 170 175
Leu Ile Arg Glu Gly Ala Thr Val Cys Gly Val Arg Val Arg Asn His
180 185 190
Leu Thr Gly Glu Thr Gln Ala Leu His Ala Pro Val Val Val Asn Ala
195 200 205
Ala Gly Ile Trp Gly Gln His Ile Ala Glu Tyr Ala Asp Leu Arg Ile
210 215 220
Arg Met Phe Pro Ala Lys Gly Ser Leu Leu Ile Met Asp His Arg Ile
225 230 235 240
Asn Gln His Val Ile Asn Arg Cys Arg Lys Pro Ser Asp Ala Asp Ile
245 250 255
Leu Val Pro Gly Asp Thr Ile Ser Leu Ile Gly Thr Thr Ser Leu Arg
260 265 270
Ile Asp Tyr Asn Glu Ile Asp Asp Asn Arg Val Thr Ala Glu Glu Val
275 280 285
Asp Ile Leu Leu Arg Glu Gly Glu Lys Leu Ala Pro Val Met Ala Lys
290 295 300
Thr Arg Ile Leu Arg Ala Tyr Ser Gly Val Arg Pro Leu Val Ala Ser
305 310 315 320
Asp Asp Asp Pro Ser Gly Arg Asn Val Ser Arg Gly Ile Val Leu Leu
325 330 335
Asp His Ala Glu Arg Asp Gly Leu Asp Gly Phe Ile Thr Ile Thr Gly
340 345 350
Gly Lys Leu Met Thr Tyr Arg Leu Met Ala Glu Trp Ala Thr Asp Ala
355 360 365
Val Cys Arg Lys Leu Gly Asn Thr Arg Pro Cys Thr Thr Ala Asp Leu
370 375 380
Ala Leu Pro Gly Ser Gln Glu Pro Ala Glu Val Thr Leu Arg Lys Val
385 390 395 400
Ile Ser Leu Pro Ala Pro Leu Arg Gly Ser Ala Val Tyr Arg His Gly
405 410 415
Asp Arg Thr Pro Ala Trp Leu Ser Glu Gly Arg Leu His Arg Ser Leu
420 425 430
Val Cys Glu Cys Glu Ala Val Thr Ala Gly Glu Val Gln Tyr Ala Val
435 440 445
Glu Asn Leu Asn Val Asn Ser Leu Leu Asp Leu Arg Arg Arg Thr Arg
450 455 460
Val Gly Met Gly Thr Cys Gln Gly Glu Leu Cys Ala Cys Arg Ala Ala
465 470 475 480
Gly Leu Leu Gln Arg Phe Asn Val Thr Thr Ser Ala Gln Ser Ile Glu
485 490 495
Gln Leu Ser Thr Phe Leu Asn Glu Arg Trp Lys Gly Val Gln Pro Ile
500 505 510
Ala Trp Gly Asp Ala Leu Arg Glu Ser Glu Phe Thr Arg Trp Val Tyr
515 520 525
Gln Gly Leu Cys Gly Leu Glu Lys Glu Gln Lys Asp Ala Leu
530 535 540

<210> 415
<211> 419
<212> PRT
<213> Escherichia coli

<400> 415
Met Arg Phe Asp Thr Val Ile Met Gly Gly Gly Leu Ala Gly Leu Leu
1 5 10 15
Cys Gly Leu Gln Leu Gln Lys His Gly Leu Arg Cys Ala Ile Val Thr
20 25 30
Arg Gly Gln Ser Ala Leu His Phe Ser Ser Gly Ser Leu Asp Leu Leu

```
        35                  40                  45
Ser His Leu Pro Asp Gly Gln Pro Val Thr Asp Ile His Ser Gly Leu
    50                  55                  60
Glu Ser Leu Arg Gln Gln Ala Pro Ala His Pro Tyr Ser Leu Leu Glu
65                  70                  75                  80
Pro Gln Arg Val Leu Asp Leu Ala Cys Gln Ala Gln Ala Leu Ile Ala
                85                  90                  95
Glu Ser Gly Ala Gln Leu Gln Gly Ser Val Glu Leu Ala His Gln Arg
            100                 105                 110
Val Thr Pro Leu Gly Thr Leu Arg Ser Thr Trp Leu Ser Ser Pro Glu
        115                 120                 125
Val Pro Val Trp Pro Leu Pro Ala Lys Lys Ile Cys Val Val Gly Ile
    130                 135                 140
Ser Gly Leu Met Asp Phe Gln Ala His Leu Ala Ala Ala Ser Leu Arg
145                 150                 155                 160
Glu Leu Gly Leu Ala Val Glu Thr Ala Glu Ile Glu Leu Pro Glu Leu
                165                 170                 175
Asp Val Leu Arg Asn Asn Ala Thr Glu Phe Arg Ala Val Asn Ile Ala
            180                 185                 190
Arg Phe Leu Asp Asn Glu Glu Asn Trp Pro Leu Leu Leu Asp Ala Leu
        195                 200                 205
Ile Pro Val Ala Asn Thr Cys Glu Met Ile Leu Met Pro Ala Cys Phe
    210                 215                 220
Gly Leu Ala Asp Asp Lys Leu Trp Arg Trp Leu Asn Glu Lys Leu Pro
225                 230                 235                 240
Cys Ser Leu Met Leu Leu Pro Thr Leu Pro Pro Ser Val Leu Gly Ile
                245                 250                 255
Arg Leu Gln Asn Gln Leu Gln Arg Gln Phe Val Arg Gln Gly Gly Val
            260                 265                 270
Trp Met Pro Gly Asp Glu Val Lys Lys Val Thr Cys Lys Asn Gly Val
        275                 280                 285
Val Asn Glu Ile Trp Thr Arg Asn His Ala Asp Ile Pro Leu Arg Pro
    290                 295                 300
Arg Phe Ala Val Leu Ala Ser Gly Ser Phe Phe Ser Gly Gly Leu Val
305                 310                 315                 320
Ala Glu Arg Asn Gly Ile Arg Glu Pro Ile Leu Gly Leu Asp Val Leu
                325                 330                 335
Gln Thr Ala Thr Arg Gly Glu Trp Tyr Lys Gly Asp Phe Phe Ala Pro
            340                 345                 350
Gln Pro Trp Gln Gln Phe Gly Val Thr Thr Asp Glu Thr Leu Arg Pro
        355                 360                 365
Ser Gln Ala Gly Gln Thr Ile Glu Asn Leu Phe Ala Ile Gly Ser Val
    370                 375                 380
Leu Gly Gly Phe Asp Pro Ile Ala Gln Gly Cys Gly Gly Gly Val Cys
385                 390                 395                 400
Ala Val Ser Ala Leu His Ala Ala Gln Gln Ile Ala Gln Arg Ala Gly
                405                 410                 415
Gly Gln Gln
```

<210> 416
<211> 396
<212> PRT
<213> Escherichia coli

<400> 416

```
Met Asn Asp Thr Ser Phe Glu Asn Cys Ile Lys Cys Thr Val Cys Thr
 1               5                  10                  15
Thr Ala Cys Pro Val Ser Arg Val Asn Pro Gly Tyr Pro Gly Pro Lys
            20                  25                  30
```

Gln Ala Gly Pro Asp Gly Glu Arg Leu Arg Leu Lys Asp Gly Ala Leu
35 40 45
Tyr Asp Glu Ala Leu Lys Tyr Cys Ile Asn Cys Lys Arg Cys Glu Val
50 55 60
Ala Cys Pro Ser Asp Val Lys Ile Gly Asp Ile Ile Gln Arg Ala Arg
65 70 75 80
Ala Lys Tyr Asp Thr Thr Arg Pro Ser Leu Arg Asn Phe Val Leu Ser
85 90 95
His Thr Asp Leu Met Gly Ser Val Ser Thr Pro Phe Ala Pro Ile Val
100 105 110
Asn Thr Ala Thr Ser Leu Lys Pro Val Arg Gln Leu Leu Asp Ala Ala
115 120 125
Leu Lys Ile Asp His Arg Arg Thr Leu Pro Lys Tyr Ser Phe Gly Thr
130 135 140
Phe Arg Arg Trp Tyr Arg Ser Val Ala Ala Gln Gln Ala Gln Tyr Lys
145 150 155 160
Asp Gln Val Ala Phe Phe His Gly Cys Phe Val Asn Tyr Asn His Pro
165 170 175
Gln Leu Gly Lys Asp Leu Ile Lys Val Leu Asn Ala Met Gly Thr Gly
180 185 190
Val Gln Leu Leu Ser Lys Glu Lys Cys Cys Gly Val Pro Leu Ile Ala
195 200 205
Asn Gly Phe Thr Asp Lys Ala Arg Lys Gln Ala Ile Thr Asn Val Glu
210 215 220
Ser Ile Arg Glu Ala Val Gly Val Lys Gly Ile Pro Val Ile Ala Thr
225 230 235 240
Ser Ser Thr Cys Thr Phe Ala Leu Arg Asp Glu Tyr Pro Glu Val Leu
245 250 255
Asn Val Asp Asn Lys Gly Leu Arg Asp His Ile Glu Leu Ala Thr Arg
260 265 270
Trp Leu Trp Arg Lys Leu Asp Glu Gly Lys Thr Leu Pro Leu Lys Pro
275 280 285
Leu Pro Leu Lys Val Val Tyr His Thr Pro Cys His Met Glu Lys Met
290 295 300
Gly Trp Thr Leu Tyr Thr Leu Glu Leu Leu Arg Asn Ile Pro Gly Leu
305 310 315 320
Glu Leu Thr Val Leu Asp Ser Gln Cys Cys Gly Ile Ala Gly Thr Tyr
325 330 335
Gly Phe Lys Lys Glu Asn Tyr Pro Thr Ser Gln Ala Ile Gly Ala Pro
340 345 350
Leu Phe Arg Gln Ile Glu Glu Ser Gly Ala Asp Leu Val Val Thr Asp
355 360 365
Cys Glu Thr Cys Lys Trp Gln Ile Glu Met Ser Thr Ser Leu Arg Cys
370 375 380
Glu His Pro Ile Thr Leu Leu Ala Gln Ala Leu Ala
385 390 395

<210> 417
<211> 302
<212> PRT
<213> Escherichia coli

<400> 417
Met Gln Leu Arg Lys Pro Ala Thr Ala Ile Leu Ala Leu Ala Leu Ser
1 5 10 15
Ala Gly Leu Ala Gln Ala Asp Asp Ala Ala Pro Ala Ala Gly Ser Thr
20 25 30
Leu Asp Lys Ile Ala Lys Asn Gly Val Ile Val Val Gly His Arg Glu
35 40 45
Ser Ser Val Pro Phe Ser Tyr Tyr Asp Asn Gln Gln Lys Val Val Gly

50 55 60
Tyr Ser Gln Asp Tyr Ser Asn Ala Ile Val Glu Ala Val Lys Lys Lys
65 70 75 80
Leu Asn Lys Pro Asp Leu Gln Val Lys Leu Ile Pro Ile Thr Ser Gln
85 90 95
Asn Arg Ile Pro Leu Leu Gln Asn Gly Thr Phe Asp Phe Glu Cys Gly
100 105 110
Ser Thr Thr Asn Asn Val Glu Arg Gln Lys Gln Ala Ala Phe Ser Asp
115 120 125
Thr Ile Phe Val Val Gly Thr Arg Leu Leu Thr Lys Lys Gly Gly Asp
130 135 140
Ile Lys Asp Phe Ala Asn Leu Lys Asp Lys Ala Val Val Val Thr Ser
145 150 155 160
Gly Thr Thr Ser Glu Val Leu Leu Asn Lys Leu Asn Glu Glu Gln Lys
165 170 175
Met Asn Met Arg Ile Ile Ser Ala Lys Asp His Gly Asp Ser Phe Arg
180 185 190
Thr Leu Glu Ser Gly Arg Ala Val Ala Phe Met Met Asp Asp Ala Leu
195 200 205
Leu Ala Gly Glu Arg Ala Lys Ala Lys Lys Pro Asp Asn Trp Glu Ile
210 215 220
Val Gly Lys Pro Gln Ser Gln Glu Ala Tyr Gly Cys Met Leu Arg Lys
225 230 235 240
Asp Asp Pro Gln Phe Lys Lys Leu Met Asp Asp Thr Ile Ala Gln Val
245 250 255
Gln Thr Ser Gly Glu Ala Glu Lys Trp Phe Asp Lys Trp Phe Lys Asn
260 265 270
Pro Ile Pro Pro Lys Asn Leu Asn Met Asn Phe Glu Leu Ser Asp Glu
275 280 285
Met Lys Ala Leu Phe Lys Glu Pro Asn Asp Lys Ala Leu Asn
290 295 300

<210> 418
<211> 328
<212> PRT
<213> Escherichia coli

<400> 418
Met Asn Asn Ser Ala Phe Thr Phe Gln Thr Leu His Pro Asp Thr Ile
1 5 10 15
Met Asp Ala Leu Phe Glu His Gly Ile Arg Val Asp Ser Gly Leu Thr
20 25 30
Pro Leu Asn Ser Tyr Glu Asn Arg Val Tyr Gln Phe Gln Asp Glu Asp
35 40 45
Arg Arg Arg Phe Val Val Lys Phe Tyr Arg Pro Glu Arg Trp Thr Ala
50 55 60
Asp Gln Ile Leu Glu Glu His Gln Phe Ala Leu Gln Leu Val Asn Asp
65 70 75 80
Glu Val Pro Val Ala Ala Pro Val Ala Phe Asn Gly Gln Thr Leu Leu
85 90 95
Asn His Gln Gly Phe Tyr Phe Ala Val Phe Pro Ser Val Gly Gly Arg
100 105 110
Gln Phe Glu Ala Asp Asn Ile Asp Gln Met Glu Ala Val Gly Arg Tyr
115 120 125
Leu Gly Arg Met His Gln Thr Gly Arg Lys Gln Leu Phe Ile His Arg
130 135 140
Pro Thr Ile Gly Leu Asn Glu Tyr Leu Ile Glu Pro Arg Lys Leu Phe
145 150 155 160
Glu Asp Ala Thr Leu Ile Pro Ser Gly Leu Lys Ala Ala Phe Leu Lys
165 170 175

Ala Thr Asp Glu Leu Ile Ala Ala Val Thr Ala His Trp Arg Glu Asp
180 185 190
Phe Thr Val Leu Arg Leu His Gly Asp Cys His Ala Gly Asn Ile Leu
195 200 205
Trp Arg Asp Gly Pro Met Phe Val Asp Leu Asp Asp Ala Arg Asn Gly
210 215 220
Pro Ala Val Gln Asp Leu Trp Met Leu Leu Asn Gly Asp Lys Ala Glu
225 230 235 240
Gln Arg Met Gln Leu Glu Thr Ile Ile Glu Ala Tyr Glu Glu Phe Ser
245 250 255
Glu Phe Asp Thr Ala Glu Ile Gly Leu Ile Glu Pro Leu Arg Ala Met
260 265 270
Arg Leu Val Tyr Tyr Leu Ala Trp Leu Met Arg Arg Trp Ala Asp Pro
275 280 285
Ala Phe Pro Lys Asn Phe Pro Trp Leu Thr Gly Glu Asp Tyr Trp Leu
290 295 300
Arg Gln Thr Ala Thr Phe Ile Glu Gln Ala Lys Val Leu Gln Glu Pro
305 310 315 320
Pro Leu Gln Leu Thr Pro Met Tyr
325

<210> 419
<211> 208
<212> PRT
<213> Escherichia coli

<400> 419
Met Lys Lys Ile Trp Leu Ala Leu Ala Gly Leu Val Leu Ala Phe Ser
1 5 10 15
Ala Ser Ala Ala Gln Tyr Glu Asp Gly Lys Gln Tyr Thr Thr Leu Glu
20 25 30
Lys Pro Val Ala Gly Ala Pro Gln Val Leu Glu Phe Phe Ser Phe Phe
35 40 45
Cys Pro His Cys Tyr Gln Phe Glu Glu Val Leu His Ile Ser Asp Asn
50 55 60
Val Lys Lys Lys Leu Pro Glu Gly Val Lys Met Thr Lys Tyr His Val
65 70 75 80
Asn Phe Met Gly Gly Asp Leu Gly Lys Asp Leu Thr Gln Ala Trp Ala
85 90 95
Val Ala Met Ala Leu Gly Val Glu Asp Lys Val Thr Val Pro Leu Phe
100 105 110
Glu Gly Val Gln Lys Thr Gln Thr Ile Arg Ser Ala Ser Asp Ile Arg
115 120 125
Asp Val Phe Ile Asn Ala Gly Ile Lys Gly Glu Glu Tyr Asp Ala Ala
130 135 140
Trp Asn Ser Phe Val Val Lys Ser Leu Val Ala Gln Gln Glu Lys Ala
145 150 155 160
Ala Ala Asp Val Gln Leu Arg Gly Val Pro Ala Met Phe Val Asn Gly
165 170 175
Lys Tyr Gln Leu Asn Pro Gln Gly Met Asp Thr Ser Asn Met Asp Val
180 185 190
Phe Val Gln Gln Tyr Ala Asp Thr Val Lys Tyr Leu Ser Glu Lys Lys
195 200 205

<210> 420
<211> 112
<212> PRT
<213> Escherichia coli

<400> 420
Met Thr Met Asn Ser Phe Glu Arg Arg Asn Lys Ile Ile Gln Leu Val
1   5   10   15
Asn Glu Gln Gly Thr Val Leu Val Gln Asp Leu Ala Gly Val Phe Ala
            20   25   30
Ala Ser Glu Ala Thr Ile Arg Ala Asp Leu Arg Phe Leu Glu Gln Lys
        35   40   45
Gly Val Val Thr Arg Phe His Gly Gly Ala Ala Lys Ile Met Ser Gly
    50   55   60
Asn Ser Glu Thr Glu Thr Gln Glu Val Gly Phe Lys Glu Arg Phe Gln
65   70   75   80
Leu Ala Ser Ala Pro Lys Asn Arg Ile Ala Gln Ala Ala Val Lys Met
                85   90   95
Ile His Glu Gly Met Thr Asp Pro Thr His Val Ile Trp Thr Gln Ala
            100   105   110

<210> 421
<211> 346
<212> PRT
<213> Escherichia coli

<400> 421
Met Lys Ser Val Val Asn Asp Thr Asp Gly Ile Val Arg Val Ala Glu
1   5   10   15
Ser Val Ile Pro Glu Ile Lys His Gln Asp Glu Val Arg Val Lys Ile
            20   25   30
Ala Ser Ser Gly Leu Cys Gly Ser Asp Leu Pro Arg Ile Phe Lys Asn
        35   40   45
Gly Ala His Tyr Tyr Pro Ile Thr Leu Gly His Glu Phe Ser Gly Tyr
    50   55   60
Ile Asp Ala Val Gly Ser Gly Val Asp Asp Leu His Pro Gly Asp Ala
65   70   75   80
Val Ala Cys Val Pro Leu Leu Pro Cys Phe Thr Cys Pro Glu Cys Leu
                85   90   95
Lys Gly Phe Tyr Ser Gln Cys Ala Lys Tyr Asp Phe Ile Gly Ser Arg
            100   105   110
Arg Asp Gly Gly Phe Ala Glu Tyr Ile Val Val Lys Arg Lys Asn Val
        115   120   125
Phe Ala Leu Pro Thr Asp Met Pro Ile Glu Asp Gly Ala Phe Ile Glu
    130   135   140
Pro Ile Thr Val Gly Leu His Ala Phe His Leu Ala Gln Gly Cys Glu
145   150   155   160
Asn Lys Asn Val Ile Ile Ile Gly Ala Gly Thr Ile Gly Leu Leu Ala
                165   170   175
Ile Gln Cys Ala Val Ala Leu Gly Ala Lys Ser Val Thr Ala Ile Asp
            180   185   190
Ile Ser Ser Glu Lys Leu Ala Leu Ala Lys Ser Phe Gly Ala Met Gln
        195   200   205
Thr Phe Asn Ser Ser Glu Met Ser Ala Pro Gln Met Gln Ser Val Leu
    210   215   220
Arg Glu Leu Arg Phe Asn Gln Leu Ile Leu Glu Thr Ala Gly Val Pro
225   230   235   240
Gln Thr Val Glu Leu Ala Val Glu Ile Ala Gly Pro His Ala Gln Leu
                245   250   255
Ala Leu Val Gly Thr Leu His Gln Asp Leu His Leu Thr Ser Ala Thr
            260   265   270
Phe Gly Lys Ile Leu Arg Lys Glu Leu Thr Val Ile Gly Ser Trp Met
        275   280   285
Asn Tyr Ser Ser Pro Trp Pro Gly Gln Glu Trp Glu Thr Ala Ser Arg
    290   295   300

Leu Leu Thr Glu Arg Lys Leu Ser Leu Glu Pro Leu Ile Ala His Arg
305 310 315 320
Gly Ser Phe Glu Ser Phe Ala Gln Ala Val Arg Asp Ile Ala Arg Asn
325 330 335
Ala Met Pro Gly Lys Val Leu Leu Ile Pro
340 345

<210> 422
<211> 451
<212> PRT
<213> Escherichia coli

<400> 422
Met Phe Ser Glu Val Met Arg Tyr Ile Leu Asp Leu Gly Pro Thr Val
1 5 10 15
Met Leu Pro Ile Val Ile Ile Ile Phe Ser Lys Ile Leu Gly Met Lys
20 25 30
Ala Gly Asp Cys Phe Lys Ala Gly Leu His Ile Gly Ile Gly Phe Val
35 40 45
Gly Ile Gly Leu Val Ile Gly Leu Met Leu Asp Ser Ile Gly Pro Ala
50 55 60
Ala Lys Ala Met Ala Glu Asn Phe Asp Leu Asn Leu His Val Val Asp
65 70 75 80
Val Gly Trp Pro Gly Ser Ser Pro Met Thr Trp Ala Ser Gln Ile Ala
85 90 95
Leu Val Ala Ile Pro Ile Ala Ile Leu Val Asn Val Ala Met Leu Leu
100 105 110
Thr Arg Met Thr Arg Val Val Asn Val Asp Ile Trp Asn Ile Trp His
115 120 125
Met Thr Phe Thr Gly Ala Leu Leu His Leu Ala Thr Gly Ser Trp Met
130 135 140
Ile Gly Met Ala Gly Val Val Ile His Ala Ala Phe Val Tyr Lys Leu
145 150 155 160
Gly Asp Trp Phe Ala Arg Asp Thr Arg Asn Phe Phe Glu Leu Glu Gly
165 170 175
Ile Ala Ile Pro His Gly Thr Ser Ala Tyr Met Gly Pro Ile Ala Val
180 185 190
Leu Val Asp Ala Ile Ile Glu Lys Ile Pro Gly Val Asn Arg Ile Lys
195 200 205
Phe Ser Ala Asp Asp Ile Gln Arg Lys Phe Gly Pro Phe Gly Glu Pro
210 215 220
Val Thr Val Gly Phe Val Met Gly Leu Ile Ile Gly Ile Leu Ala Gly
225 230 235 240
Tyr Asp Val Lys Gly Val Leu Gln Leu Ala Val Lys Thr Ala Ala Val
245 250 255
Met Leu Leu Met Pro Arg Val Ile Lys Pro Ile Met Asp Gly Leu Thr
260 265 270
Pro Ile Ala Lys Gln Ala Arg Ser Arg Leu Gln Ala Lys Phe Gly Gly
275 280 285
Gln Glu Phe Leu Ile Gly Leu Asp Pro Ala Leu Leu Leu Gly His Thr
290 295 300
Ala Val Val Ser Ala Ser Leu Ile Phe Ile Pro Leu Thr Ile Leu Ile
305 310 315 320
Ala Val Cys Val Pro Gly Asn Gln Val Leu Pro Phe Gly Asp Leu Ala
325 330 335
Thr Ile Gly Phe Phe Val Ala Met Ala Val Ala Val His Arg Gly Asn
340 345 350
Leu Phe Arg Thr Leu Ile Ser Gly Val Ile Ile Met Ser Ile Thr Leu
355 360 365
Trp Ile Ala Thr Gln Thr Ile Gly Leu His Thr Gln Leu Ala Ala Asn

```
    370                 375                 380
Ala Gly Ala Leu Lys Ala Gly Gly Met Val Ala Ser Met Asp Gln Gly
385                 390                 395                 400
Gly Ser Pro Ile Thr Trp Leu Leu Ile Gln Val Phe Ser Pro Gln Asn
                405                 410                 415
Ile Pro Gly Phe Ile Ile Ile Gly Ala Ile Tyr Leu Thr Gly Ile Phe
            420                 425                 430
Met Thr Trp Arg Arg Ala Arg Gly Phe Ile Lys Gln Glu Lys Val Val
        435                 440                 445
Leu Ala Glu
    450
```

<210> 423
<211> 94
<212> PRT
<213> Escherichia coli

<400> 423

```
Met Lys Arg Lys Ile Ile Val Ala Cys Gly Gly Ala Val Ala Thr Ser
 1               5                  10                  15
Thr Met Ala Ala Glu Glu Ile Lys Glu Leu Cys Gln Asn His Asn Ile
            20                  25                  30
Pro Val Glu Leu Ile Gln Cys Arg Val Asn Glu Ile Glu Thr Tyr Met
        35                  40                  45
Asp Gly Val His Leu Ile Cys Thr Thr Ala Lys Val Asp Arg Ser Phe
    50                  55                  60
Gly Asp Ile Pro Leu Val His Gly Met Pro Phe Ile Ser Gly Ile Gly
65                  70                  75                  80
Ile Glu Ala Leu Gln Asn Lys Ile Leu Thr Ile Leu Gln Gly
                85                  90
```

<210> 424
<211> 150
<212> PRT
<213> Escherichia coli

<400> 424

```
Met Thr Asn Leu Phe Val Arg Ser Gly Ile Ser Phe Val Asp Arg Ser
 1               5                  10                  15
Glu Val Leu Thr His Ile Gly Asn Glu Met Leu Ala Lys Gly Val Val
            20                  25                  30
His Asp Thr Trp Pro Gln Ala Leu Ile Ala Arg Glu Ala Glu Phe Pro
        35                  40                  45
Thr Gly Ile Met Leu Glu Gln His Ala Ile Ala Ile Pro His Cys Glu
    50                  55                  60
Ala Ile His Ala Lys Ser Ser Ala Ile Tyr Leu Leu Arg Pro Thr Asn
65                  70                  75                  80
Lys Val His Phe Gln Gln Ala Asp Asp Asp Asn Asp Val Ala Val Ser
                85                  90                  95
Leu Val Ile Ala Leu Ile Val Glu Asn Pro Gln Gln Gln Leu Lys Leu
            100                 105                 110
Leu Arg Cys Leu Phe Gly Lys Leu Gln Gln Pro Asp Ile Val Glu Thr
        115                 120                 125
Leu Ile Thr Leu Pro Glu Thr Gln Leu Lys Glu Tyr Phe Thr Lys Tyr
    130                 135                 140
Val Leu Asp Ser Asp Glu
145                 150
```

<210> 425
<211> 420
<212> PRT
<213> Escherichia coli

<400> 425
Met Lys Thr Leu Ile Ala Arg His Lys Ala Gly Glu His Ile Gly Ile
1 5 10 15
Cys Ser Val Cys Ser Ala His Pro Leu Val Ile Glu Ala Ala Leu Ala
20 25 30
Phe Asp Arg Asn Ser Thr Arg Lys Val Leu Ile Glu Ala Thr Ser Asn
35 40 45
Gln Val Asn Gln Phe Gly Gly Tyr Thr Gly Met Thr Pro Ala Asp Phe
50 55 60
Arg Glu Phe Val Phe Thr Ile Ala Asp Lys Val Gly Phe Ala Arg Glu
65 70 75 80
Arg Ile Ile Leu Gly Gly Asp His Leu Gly Pro Asn Cys Trp Gln Gln
85 90 95
Glu Asn Ala Asp Ala Ala Met Glu Lys Ser Val Glu Leu Val Lys Glu
100 105 110
Tyr Val Arg Ala Gly Phe Ser Lys Ile His Leu Asp Ala Ser Met Ser
115 120 125
Cys Ala Gly Asp Pro Ile Pro Leu Ala Pro Glu Thr Val Ala Glu Arg
130 135 140
Ala Ala Val Leu Cys Phe Ala Ala Glu Ser Val Ala Thr Asp Cys Gln
145 150 155 160
Arg Glu Gln Leu Ser Tyr Val Ile Gly Thr Glu Val Pro Val Pro Gly
165 170 175
Gly Glu Ala Ser Ala Ile Gln Ser Val His Ile Thr His Val Glu Asp
180 185 190
Ala Ala Asn Thr Leu Arg Thr His Gln Lys Ala Phe Ile Ala Arg Gly
195 200 205
Leu Thr Glu Ala Leu Thr Arg Val Ile Ala Ile Val Val Gln Pro Gly
210 215 220
Val Glu Phe Asp His Ser Asn Ile Ile His Tyr Gln Pro Gln Glu Ala
225 230 235 240
Gln Pro Leu Ala Gln Trp Ile Glu Asn Thr Arg Met Val Tyr Glu Ala
245 250 255
His Ser Thr Asp Tyr Gln Thr Arg Thr Ala Tyr Trp Glu Leu Val Arg
260 265 270
Asp His Phe Ala Ile Leu Lys Val Gly Pro Ala Leu Thr Phe Ala Leu
275 280 285
Arg Glu Ala Ile Phe Ala Leu Ala Gln Ile Glu Gln Glu Leu Ile Ala
290 295 300
Pro Glu Asn Arg Ser Gly Cys Leu Ala Val Ile Glu Glu Val Met Leu
305 310 315 320
Asp Glu Pro Gln Tyr Trp Lys Lys Tyr Tyr Arg Thr Gly Phe Asn Asp
325 330 335
Ser Leu Leu Asp Ile Arg Tyr Ser Leu Ser Asp Arg Ile Arg Tyr Tyr
340 345 350
Trp Pro His Ser Arg Ile Lys Asn Ser Val Glu Thr Met Met Val Asn
355 360 365
Leu Glu Gly Val Asp Ile Pro Leu Gly Met Ile Ser Gln Tyr Leu Pro
370 375 380
Lys Gln Phe Glu Arg Ile Gln Ser Gly Glu Leu Ser Ala Ile Pro His
385 390 395 400
Gln Leu Ile Met Asp Lys Ile Tyr Asp Val Leu Arg Ala Tyr Arg Tyr
405 410 415
Gly Cys Ala Glu
420

```
<210> 426
<211> 286
<212> PRT
<213> Escherichia coli

<400> 426
Met Lys Met Tyr Val Val Ser Thr Lys Gln Met Leu Asn Asn Ala Gln
 1               5                  10                  15
Arg Gly Gly Tyr Ala Val Pro Ala Phe Asn Ile His Asn Leu Glu Thr
            20                  25                  30
Met Gln Val Val Val Glu Thr Ala Ala Asn Leu His Ala Pro Val Ile
        35                  40                  45
Ile Ala Gly Thr Pro Gly Thr Phe Thr His Ala Gly Thr Glu Asn Leu
    50                  55                  60
Leu Ala Leu Val Ser Ala Met Ala Lys Gln Tyr His His Pro Leu Ala
65                  70                  75                  80
Ile His Leu Asp His His Thr Lys Phe Asp Asp Ile Ala Gln Lys Val
                85                  90                  95
Arg Ser Gly Val Arg Ser Val Met Ile Asp Ala Ser His Leu Pro Phe
            100                 105                 110
Ala Gln Asn Ile Ser Arg Val Lys Glu Val Val Asp Phe Cys His Arg
        115                 120                 125
Phe Asp Val Ser Val Glu Ala Glu Leu Gly Gln Leu Gly Gly Gln Glu
    130                 135                 140
Asp Asp Val Gln Val Asn Glu Ala Asp Ala Leu Tyr Thr Asn Pro Ala
145                 150                 155                 160
Gln Ala Arg Glu Phe Ala Glu Ala Thr Gly Ile Asp Ser Leu Ala Val
                165                 170                 175
Ala Ile Gly Thr Ala His Gly Met Tyr Ala Ser Ala Pro Ala Leu Asp
            180                 185                 190
Phe Ser Arg Leu Glu Asn Ile Arg Gln Trp Val Asn Leu Pro Leu Val
        195                 200                 205
Leu His Gly Ala Ser Gly Leu Ser Thr Lys Asp Ile Gln Gln Thr Ile
    210                 215                 220
Lys Leu Gly Ile Cys Lys Ile Asn Val Ala Thr Glu Leu Lys Asn Ala
225                 230                 235                 240
Phe Ser Gln Ala Leu Lys Asn Tyr Leu Thr Glu His Pro Glu Ala Thr
                245                 250                 255
Asp Pro Arg Asp Tyr Leu Gln Ser Ala Lys Ser Ala Met Arg Asp Val
            260                 265                 270
Val Ser Lys Val Ile Ala Asp Cys Gly Cys Glu Gly Arg Ala
        275                 280                 285


<210> 427
<211> 157
<212> PRT
<213> Escherichia coli

<400> 427
Met Ser Gln Asn Asp Ile Ile Ile Arg Thr His Tyr Lys Ser Pro His
 1               5                  10                  15
Arg Leu His Ile Asp Ser Asp Ile Pro Thr Pro Ser Ser Glu Pro Ile
            20                  25                  30
Asn Gln Phe Ala Arg Gln Leu Ile Thr Leu Leu Asp Thr Ser Asp Leu
        35                  40                  45
Ser Ser Met Leu Ser Tyr Cys Val Thr Gln Glu Phe Thr Ala Asn Cys
    50                  55                  60
Arg Lys Ile Ser Gln Asn Cys Tyr Ser Thr Ala Leu Phe Thr Ile Asn
65                  70                  75                  80
```

Phe Ala Thr Ser Pro Ile His Thr Glu Asn Ile Leu Ile Thr Leu His
85 90 95
Tyr Lys Lys Glu Ile Ile Ser Leu Leu Leu Glu Thr Thr Pro Ile Lys
100 105 110
Ala Asn His Leu Arg Ser Ile Leu Asp Tyr Ile Glu Gln Glu Gln Leu
115 120 125
Thr Ala Glu Asp Arg Asn His Cys Met Lys Leu Ser Lys Lys Ile His
130 135 140
Arg Glu Lys Asn Tyr Thr Pro Asn Ser Lys Ser Gln Trp
145 150 155

<210> 428
<211> 471
<212> PRT
<213> Escherichia coli

<400> 428
Met Gln Ser Pro Ser Asp Ala Ile Phe Cys Arg His Leu Ser Leu Gln
1 5 10 15
Tyr Ala Leu Asp Ser Leu Arg Asn Gly Lys Gly Lys Val Asn Leu Ile
20 25 30
Lys His Tyr Ser Ser Val Glu Ser Ile Gln Gln His Val Pro Leu Val
35 40 45
Arg Asp Ala Glu Phe Arg Ala Leu Leu Arg His Pro Pro Ala Gly Ser
50 55 60
Arg Val Ile Ala Ser Lys Asp Phe Gly Phe Ala Leu Asp Ile Phe Phe
65 70 75 80
Cys Arg Met Met Ala Asn Asn Val Ser His Met Ser Ala Ile Leu Tyr
85 90 95
Ile Asp Asn His Thr Leu Ser Val Arg Leu Arg Ile Lys Gln Ser Val
100 105 110
Tyr Gly Gln Leu Asn Tyr Val Val Ser Val Tyr Asp Pro Asn Asp Thr
115 120 125
Asn Val Ala Val Arg Asp Thr His Arg Thr Ala Arg Gly Phe Leu Ser
130 135 140
Leu Asp Lys Phe Ile Ser Ser Gly Pro Asp Ala Gln Thr Trp Ala Asp
145 150 155 160
Arg Tyr Val Arg Asn Cys Ala Ile Ala Ile Leu Pro Leu Leu Pro Val
165 170 175
Gly Val Pro Gly Ala Ile Phe Ala Gly Ile Ala Ser Arg Met Pro Phe
180 185 190
Ala Pro Ile His Pro Ser Ala Met Leu Leu Ile Met Ala Thr Gly Gln
195 200 205
Ser Gln Gln Leu Ile Thr Leu Phe Lys Gln Leu Pro Ile Leu Pro Glu
210 215 220
Lys Glu Ile Ile Glu Ile Ile Thr Ala Gln Asn Ser Val Gly Thr Pro
225 230 235 240
Ala Leu Phe Leu Ala Met Met Asn Gly His Thr Asp Asn Val Lys Ile
245 250 255
Phe Met Gln Glu Ile Gln Ser Leu Val Asp Asn His Ile Ile His Glu
260 265 270
Asp Asn Leu Val Lys Leu Leu Gln Thr Lys Ser Ala Asn Glu Thr Pro
275 280 285
Gly Leu Tyr Ile Ser Met Leu Tyr Gly Phe Asp Glu Ile Ile Asp Ile
290 295 300
Phe Leu Asn Ala Leu Thr Thr Pro Ile Ala Gln Glu Leu Leu Asn Lys
305 310 315 320
Lys Leu Val Met Ser Ile Leu Ala Met Lys Ile His Asp Gly Glu Pro
325 330 335
Gly Leu Tyr Ala Ala Met Glu Asn Asn His Pro Leu Cys Val Thr Arg

340 345 350
Phe Leu Ser Lys Ile Asn Gly Ile Ala Phe Lys Tyr Lys Leu Ser Lys
355 360 365
Ala Asn Ile Met Asp Leu Leu Lys Gly Ala Thr Ala Gln Gly Thr Pro
370 375 380
Ala Leu Tyr Ile Ala Met Ser Lys Gly Asn Glu Asp Val Val Leu Ser
385 390 395 400
Tyr Ile Ser Thr Leu Gly Ala Phe Ala Lys Lys His Ser Phe Ser Gln
405 410 415
His Gln Leu Phe Thr Leu Leu Ala Ala Lys Asn His Asp Asn Met Ser
420 425 430
Ala Val His Ile Ala Ile His His Lys His Tyr Lys Thr Val Glu Thr
435 440 445
Tyr Tyr Ala Ala Ile Asn Ala Ile Ser Gln Ser Leu Ser Phe Ser Ala
450 455 460
Asp Glu Ile Lys Thr Tyr Leu
465 470

<210> 429
<211> 128
<212> PRT
<213> Escherichia coli

<400> 429
Met Pro Ser Gly Leu Phe Met Asp Leu Leu Pro Phe Leu Leu Asp Ala
1 5 10 15
Asn Leu Ser Ala Thr Asn Pro Pro Ala Ile Pro His Trp Trp Lys Arg
20 25 30
Gln Pro Leu Ile Pro Asn Leu Leu Ser Gln Glu Leu Lys Asn Tyr Leu
35 40 45
Lys Leu Asn Val Lys Glu Lys Asn Ile Gln Ile Ala Asp Gln Val Ile
50 55 60
Ile Asp Glu Thr Ala Gly Glu Val Val Ile Gly Ala Asn Thr Arg Ile
65 70 75 80
Cys His Gly Ala Val Ile Gln Gly Pro Val Val Ile Gly Ala Asn Cys
85 90 95
Leu Ile Gly Asn Trp Ile Cys Pro Tyr Ile Ser Arg His Leu Leu Ser
100 105 110
Leu Asn Pro Leu Gln Ala Arg Cys Arg Arg Tyr Ser Arg Gly Glu Arg
115 120 125

<210> 430
<211> 398
<212> PRT
<213> Escherichia coli

<400> 430
Met Lys Thr Trp Ile Phe Ile Cys Met Ser Ile Ala Met Leu Leu Trp
1 5 10 15
Phe Leu Ser Thr Leu Arg Arg Lys Pro Ser Gln Lys Lys Gly Cys Ile
20 25 30
Asp Ala Ile Ile Pro Ala Tyr Asn Glu Gly Pro Cys Leu Ala Gln Ser
35 40 45
Leu Asp Asn Leu Leu Arg Asn Pro Tyr Phe Cys Arg Val Ile Cys Val
50 55 60
Asn Asp Gly Ser Thr Asp Asn Thr Glu Ala Val Met Ala Glu Val Lys
65 70 75 80
Arg Lys Trp Gly Asp Arg Phe Val Ala Val Thr Gln Lys Asn Thr Gly
85 90 95

Lys Gly Gly Ala Leu Met Asn Gly Leu Asn Tyr Ala Thr Cys Asp Gln
100 105 110
Val Phe Leu Ser Asp Ala Asp Thr Tyr Val Pro Pro Asp Gln Asp Gly
115 120 125
Met Gly Tyr Met Leu Ala Glu Ile Glu Arg Gly Ala Asp Ala Val Gly
130 135 140
Gly Ile Pro Ser Thr Ala Leu Lys Gly Ala Gly Leu Leu Pro His Ile
145 150 155 160
Arg Ala Thr Val Lys Leu Pro Met Ile Val Met Lys Arg Thr Leu Gln
165 170 175
Gln Leu Leu Gly Gly Ala Pro Phe Ile Ile Ser Gly Ala Cys Gly Met
180 185 190
Phe Arg Thr Asp Val Leu Arg Lys Phe Gly Phe Ser Asp Arg Thr Lys
195 200 205
Val Glu Asp Leu Asp Leu Thr Trp Thr Leu Val Ala Asn Gly Tyr Arg
210 215 220
Ile Arg Gln Ala Asn Arg Cys Ile Val Tyr Pro Gln Glu Cys Asn Ser
225 230 235 240
Pro Arg Glu Glu Trp Arg Arg Trp Arg Arg Trp Ile Val Gly Tyr Ala
245 250 255
Val Cys Met Arg Leu His Lys Arg Leu Leu Phe Ser Arg Phe Gly Ile
260 265 270
Phe Ser Ile Phe Pro Met Leu Leu Val Val Leu Tyr Gly Val Gly Ile
275 280 285
Tyr Leu Thr Thr Trp Phe Asn Glu Phe Ile Thr Thr Gly Pro His Gly
290 295 300
Val Val Leu Ala Met Phe Pro Leu Ile Trp Val Gly Val Val Cys Val
305 310 315 320
Ile Gly Ala Phe Ser Ala Trp Phe His Arg Cys Trp Leu Leu Val Pro
325 330 335
Leu Ala Pro Leu Ser Val Val Tyr Val Leu Leu Ala Tyr Ala Ile Trp
340 345 350
Ile Ile Tyr Gly Leu Ile Ala Phe Phe Thr Gly Arg Glu Pro Gln Arg
355 360 365
Asp Lys Pro Thr Arg Tyr Ser Ala Leu Val Glu Ala Ser Thr Ala Tyr
370 375 380
Ser Gln Pro Ser Val Thr Gly Thr Glu Lys Leu Ser Glu Ala
385 390 395

<210> 431
<211> 552
<212> PRT
<213> Escherichia coli

<400> 431
Met Ile Leu Glu Arg Val Glu Ile Val Gly Phe Arg Gly Ile Asn Arg
1 5 10 15
Leu Ser Leu Met Leu Glu Gln Asn Asn Val Leu Ile Gly Glu Asn Ala
20 25 30
Trp Gly Lys Ser Ser Leu Leu Asp Ala Leu Thr Leu Leu Leu Ser Pro
35 40 45
Glu Ser Asp Leu Tyr His Phe Glu Arg Asp Asp Phe Trp Phe Pro Pro
50 55 60
Gly Asp Ile Asn Gly Arg Glu His His Leu His Ile Ile Leu Thr Phe
65 70 75 80
Arg Glu Ser Leu Pro Gly Arg His Arg Val Arg Arg Tyr Arg Pro Leu
85 90 95
Glu Ala Cys Trp Thr Pro Cys Thr Asp Gly Tyr His Arg Ile Phe Tyr
100 105 110
Arg Leu Glu Gly Glu Ser Ala Glu Asp Gly Ser Val Met Thr Leu Arg

115 120 125
Ser Phe Leu Asp Lys Asp Gly His Pro Ile Asp Val Glu Asp Ile Asn
130 135 140
Asp Gln Ala Arg His Leu Val Arg Leu Met Pro Val Leu Arg Leu Arg
145 150 155 160
Asp Ala Arg Phe Met Arg Arg Ile Arg Asn Gly Thr Val Pro Asn Val
165 170 175
Pro Asn Val Glu Val Thr Ala Arg Gln Leu Asp Phe Leu Ala Arg Glu
180 185 190
Leu Ser Ser His Pro Gln Asn Leu Ser Asp Gly Gln Ile Arg Gln Gly
195 200 205
Leu Ser Ala Met Val Gln Leu Leu Glu His Tyr Phe Ser Glu Gln Gly
210 215 220
Ala Gly Gln Ala Arg Tyr Arg Leu Met Arg Arg Arg Ala Ser Asn Glu
225 230 235 240
Gln Arg Ser Trp Arg Tyr Leu Asp Ile Ile Asn Arg Met Ile Asp Arg
245 250 255
Pro Gly Gly Arg Ser Tyr Arg Val Ile Leu Leu Gly Leu Phe Ala Thr
260 265 270
Leu Leu Gln Ala Lys Gly Thr Leu Arg Leu Asp Lys Asp Ala Arg Pro
275 280 285
Leu Leu Leu Ile Glu Asp Pro Glu Thr Arg Leu His Pro Ile Met Leu
290 295 300
Ser Val Ala Trp His Leu Leu Asn Leu Leu Pro Leu Gln Arg Ile Ala
305 310 315 320
Thr Thr Asn Ser Gly Glu Leu Leu Ser Leu Thr Pro Val Glu His Val
325 330 335
Cys Arg Leu Val Arg Glu Ser Ser Arg Val Ala Ala Trp Arg Leu Gly
340 345 350
Pro Ser Gly Leu Ser Thr Glu Asp Ser Arg Arg Ile Ser Phe His Ile
355 360 365
Arg Phe Asn Arg Pro Ser Ser Leu Phe Ala Arg Cys Trp Leu Leu Val
370 375 380
Glu Gly Glu Thr Glu Thr Trp Val Ile Asn Glu Leu Ala Arg Gln Cys
385 390 395 400
Gly His His Phe Asp Ala Glu Gly Ile Lys Val Ile Glu Phe Ala Gln
405 410 415
Ser Gly Leu Lys Pro Leu Val Lys Phe Ala Arg Arg Met Gly Ile Glu
420 425 430
Trp His Val Leu Val Asp Gly Asp Glu Ala Gly Lys Lys Tyr Ala Ala
435 440 445
Thr Val Arg Ser Leu Leu Asn Asn Asp Arg Glu Ala Glu Arg Glu His
450 455 460
Leu Thr Ala Leu Pro Ala Leu Asp Met Glu His Phe Met Tyr Arg Gln
465 470 475 480
Gly Phe Ser Asp Val Phe His Arg Met Ala Gln Ile Pro Glu Asn Val
485 490 495
Pro Met Asn Leu Arg Lys Ile Ile Ser Lys Ala Ile His Arg Ser Ser
500 505 510
Lys Pro Asp Leu Ala Ile Glu Val Ala Met Glu Ala Gly Arg Arg Gly
515 520 525
Val Asp Ser Val Pro Thr Leu Leu Lys Lys Met Phe Ser Arg Val Leu
530 535 540
Trp Leu Ala Arg Gly Arg Ala Asp
545 550

<210> 432
<211> 352
<212> PRT
<213> Escherichia coli

<400> 432
Met Leu Pro Ser Ile Ser Ile Asn Asn Thr Ser Ala Ala Tyr Pro Glu
1 5 10 15
Ser Ile Asn Glu Asn Asn Asn Asp Glu Val Asn Gly Leu Val Gln Glu
20 25 30
Phe Lys Asn Leu Phe Asn Gly Lys Glu Gly Ile Ser Thr Cys Ile Lys
35 40 45
His Leu Leu Glu Leu Ile Lys Asn Ala Ile Arg Val Asn Asp Asp Pro
50 55 60
Tyr Arg Phe Asn Ile Asn Asn Ser Ser Val Thr Tyr Ile Asp Ile Asp
65 70 75 80
Ser Asn Asp Thr Asp His Ile Thr Ile Gly Ile Asp Asn Gln Glu Pro
85 90 95
Ile Glu Leu Pro Ala Asn Tyr Lys Asp Lys Glu Leu Val Arg Thr Ile
100 105 110
Ile Asn Asp Asn Ile Val Glu Lys Thr His Asp Ile Asn Asn Lys Glu
115 120 125
Met Ile Phe Ser Ala Leu Lys Glu Ile Tyr Asp Gly Asp Pro Gly Phe
130 135 140
Ile Phe Asp Lys Ile Ser His Lys Leu Arg His Thr Val Thr Glu Phe
145 150 155 160
Asp Glu Ser Gly Lys Ser Glu Pro Thr Asp Leu Phe Thr Trp Tyr Gly
165 170 175
Lys Asp Lys Lys Gly Asp Ser Leu Ala Ile Val Ile Lys Asn Lys Asn
180 185 190
Gly Asn Asp Tyr Leu Ser Leu Gly Tyr Tyr Asp Gln Asp Asp Tyr His
195 200 205
Ile Gln Arg Gly Ile Arg Ile Asn Gly Asp Ser Leu Thr Gln Tyr Cys
210 215 220
Ser Glu Asn Ala Arg Ser Ala Ser Ala Trp Phe Glu Ser Ser Lys Ala
225 230 235 240
Ile Met Ala Glu Ser Phe Ala Thr Gly Ser Asp His Gln Val Val Asn
245 250 255
Glu Leu Asn Gly Glu Arg Leu Arg Glu Pro Asn Asp Val Phe Lys Arg
260 265 270
Tyr Gly Arg Ala Ile Arg Tyr Asp Phe Gln Val Asp Asp Ala Lys Tyr
275 280 285
Lys Cys Asp His Leu Lys Glu Ile Val Ser Thr Leu Val Gly Asn Lys
290 295 300
Ile Asn Val Gly His Ser Gln Lys Ile Tyr Lys His Phe Lys Asp Leu
305 310 315 320
Glu Gly Lys Ile Glu Glu Arg Leu Gln Asn Arg Gln Ala Glu Tyr Gln
325 330 335
Asn Glu Ile Asn Gln Pro Ser Ala Pro Gly Val Asn Phe Asp Asp Ile
340 345 350

<210> 433
<211> 375
<212> PRT
<213> Escherichia coli

<400> 433
Met Met Arg His Leu Arg Asn Ile Phe Asn Leu Gly Ile Lys Glu Leu
1 5 10 15
Arg Ser Leu Leu Gly Asp Lys Ala Met Leu Thr Leu Ile Val Phe Ser
20 25 30
Phe Thr Val Ser Val Tyr Ser Ser Ala Thr Val Thr Pro Gly Ser Leu
35 40 45
Asn Leu Ala Pro Ile Ala Ile Ala Asp Met Asp Gln Ser Gln Leu Ser

```
    50                  55                  60
Asn Arg Ile Val Asn Ser Phe Tyr Arg Pro Trp Phe Leu Pro Pro Glu
65                  70                  75                  80
Met Ile Thr Ala Asp Glu Met Asp Ala Gly Leu Asp Ala Gly Arg Tyr
                85                  90                  95
Thr Phe Ala Ile Asn Ile Pro Pro Asn Phe Gln Arg Asp Val Leu Ala
            100                 105                 110
Gly Arg Gln Pro Asp Ile Gln Val Asn Val Asp Ala Thr Arg Met Ser
        115                 120                 125
Gln Ala Phe Thr Gly Asn Gly Tyr Ile Gln Asn Ile Ile Asn Gly Glu
    130                 135                 140
Val Asn Ser Phe Val Ala Arg Tyr Arg Asp Asn Ser Glu Pro Leu Val
145                 150                 155                 160
Ser Leu Glu Thr Arg Met Arg Phe Asn Pro Asn Leu Asp Pro Ala Trp
                165                 170                 175
Phe Gly Gly Val Met Ala Ile Ile Asn Asn Ile Thr Met Leu Ala Ile
            180                 185                 190
Val Leu Thr Gly Ser Ala Leu Ile Arg Glu Arg Glu His Gly Thr Val
        195                 200                 205
Glu His Leu Leu Val Met Pro Ile Thr Pro Phe Glu Ile Met Met Ala
    210                 215                 220
Lys Ile Trp Ser Met Gly Leu Val Val Leu Val Val Ser Gly Leu Ser
225                 230                 235                 240
Leu Val Leu Met Val Lys Gly Val Leu Gly Val Pro Ile Glu Gly Ser
                245                 250                 255
Ile Pro Leu Phe Met Leu Gly Val Ala Leu Ser Leu Phe Ala Thr Thr
            260                 265                 270
Ser Ile Gly Ile Phe Met Gly Thr Ile Ala Arg Ser Met Pro Gln Leu
        275                 280                 285
Gly Leu Leu Val Ile Leu Val Leu Leu Pro Leu Gln Met Leu Ser Gly
    290                 295                 300
Gly Ser Thr Pro Arg Glu Ser Met Pro Gln Met Val Gln Asp Ile Met
305                 310                 315                 320
Leu Thr Met Pro Thr Thr His Phe Val Ser Leu Ala Gln Ala Ile Leu
                325                 330                 335
Tyr Arg Gly Ala Gly Phe Glu Ile Val Trp Pro Gln Phe Leu Thr Leu
            340                 345                 350
Met Ala Ile Gly Gly Ala Phe Phe Thr Ile Ala Leu Leu Arg Phe Arg
        355                 360                 365
Lys Thr Ile Gly Thr Met Ala
    370                 375
```

```
<210> 434
<211> 894
<212> PRT
<213> Escherichia coli

<400> 434
Met Ser Gln His Tyr Gly Lys Thr Val Ala Leu Asn Asn Ile Thr Leu
 1               5                  10                  15
Asp Ile Pro Ala Arg Cys Met Val Gly Leu Ile Gly Pro Asp Gly Val
            20                  25                  30
Gly Lys Ser Ser Leu Leu Ser Leu Ile Ser Gly Ala Arg Val Ile Glu
        35                  40                  45
Gln Gly Asn Val Met Val Leu Gly Gly Asp Met Arg Asp Pro Lys His
    50                  55                  60
Arg Arg Asp Val Cys Pro Arg Ile Ala Trp Met Pro Gln Gly Leu Gly
65                  70                  75                  80
Lys Asn Leu Tyr His Thr Leu Ser Val Tyr Glu Asn Val Asp Phe Phe
                85                  90                  95
```

Ala Arg Leu Phe Gly His Asp Lys Ala Glu Arg Glu Val Arg Ile Asn
100 105 110
Glu Leu Leu Thr Ser Thr Gly Leu Ala Pro Phe Arg Asp Arg Pro Ala
115 120 125
Gly Lys Leu Ser Gly Gly Met Lys Gln Lys Leu Gly Leu Cys Cys Ala
130 135 140
Leu Ile His Asp Pro Glu Leu Leu Ile Leu Asp Glu Pro Thr Thr Gly
145 150 155 160
Val Asp Pro Leu Ser Arg Ser Gln Phe Trp Asp Leu Ile Asp Ser Ile
165 170 175
Arg Gln Arg Gln Ser Asn Met Ser Val Leu Val Ala Thr Ala Tyr Met
180 185 190
Glu Glu Ala Glu Arg Phe Asp Trp Leu Val Ala Met Asn Ala Gly Glu
195 200 205
Val Leu Ala Thr Gly Ser Ala Glu Glu Leu Arg Gln Gln Thr Gln Ser
210 215 220
Ala Thr Leu Glu Glu Ala Phe Ile Asn Leu Leu Pro Gln Ala Gln Arg
225 230 235 240
Gln Ala His Gln Ala Val Val Ile Pro Pro Tyr Gln Pro Glu Asn Ala
245 250 255
Glu Ile Ala Ile Glu Ala Arg Asp Leu Thr Met Arg Phe Gly Ser Phe
260 265 270
Val Ala Val Asp His Val Asn Phe Arg Ile Pro Arg Gly Glu Ile Phe
275 280 285
Gly Phe Leu Gly Ser Asn Gly Cys Gly Lys Ser Thr Thr Met Lys Met
290 295 300
Leu Thr Gly Leu Leu Pro Ala Ser Glu Gly Glu Ala Trp Leu Phe Gly
305 310 315 320
Gln Pro Val Asp Pro Lys Asp Ile Asp Thr Arg Arg Arg Val Gly Tyr
325 330 335
Met Ser Gln Ala Phe Ser Leu Tyr Asn Glu Leu Thr Val Arg Gln Asn
340 345 350
Leu Glu Leu His Ala Arg Leu Phe His Ile Pro Glu Ala Glu Ile Pro
355 360 365
Ala Arg Val Ala Glu Met Ser Glu Arg Phe Lys Leu Asn Asp Val Glu
370 375 380
Asp Ile Leu Pro Glu Ser Leu Pro Leu Gly Ile Arg Gln Arg Leu Ser
385 390 395 400
Leu Ala Val Ala Val Ile His Arg Pro Glu Met Leu Ile Leu Asp Glu
405 410 415
Pro Thr Ser Gly Val Asp Pro Val Ala Arg Asp Met Phe Trp Gln Leu
420 425 430
Met Val Asp Leu Ser Arg Gln Asp Lys Val Thr Ile Phe Ile Ser Thr
435 440 445
His Phe Met Asn Glu Ala Glu Arg Cys Asp Arg Ile Ser Leu Met His
450 455 460
Ala Gly Lys Val Leu Ala Ser Gly Thr Pro Gln Glu Leu Val Glu Lys
465 470 475 480
Arg Gly Ala Ala Ser Leu Glu Glu Ala Phe Ile Ala Tyr Leu Gln Glu
485 490 495
Ala Ala Gly Gln Ser Asn Glu Ala Glu Ala Pro Pro Val Val His Asp
500 505 510
Thr Thr His Ala Pro Arg Gln Gly Phe Ser Leu Arg Arg Leu Phe Ser
515 520 525
Tyr Ser Arg Arg Glu Ala Leu Glu Leu Arg Arg Asp Pro Val Arg Ser
530 535 540
Thr Leu Ala Leu Met Gly Thr Val Ile Leu Met Leu Ile Met Gly Tyr
545 550 555 560
Gly Ile Ser Met Asp Val Glu Asn Leu Arg Phe Ala Val Leu Asp Arg
565 570 575
Asp Gln Thr Val Ser Ser Gln Ala Trp Thr Leu Asn Leu Ser Gly Ser

```
            580                 585                 590
Arg Tyr Phe Ile Glu Gln Pro Pro Leu Thr Ser Tyr Asp Glu Leu Asp
        595                 600                 605
Arg Arg Met Arg Ala Gly Asp Ile Thr Val Ala Ile Glu Ile Pro Pro
    610                 615                 620
Asn Phe Gly Arg Asp Ile Ala Arg Gly Thr Pro Val Glu Leu Gly Val
625                 630                 635                 640
Trp Ile Asp Gly Ala Met Pro Ser Arg Ala Glu Thr Val Lys Gly Tyr
                645                 650                 655
Val Gln Ala Met His Gln Ser Trp Leu Gln Asp Val Ala Ser Arg Gln
            660                 665                 670
Ser Thr Pro Ala Ser Gln Ser Gly Leu Met Asn Ile Glu Thr Arg Tyr
        675                 680                 685
Arg Tyr Asn Pro Asp Val Lys Ser Leu Pro Ala Ile Val Pro Ala Val
    690                 695                 700
Ile Pro Leu Leu Leu Met Met Ile Pro Ser Met Leu Ser Ala Leu Ser
705                 710                 715                 720
Val Val Arg Glu Lys Glu Leu Gly Ser Ile Ile Asn Leu Tyr Val Thr
                725                 730                 735
Pro Thr Thr Arg Ser Glu Phe Leu Leu Gly Lys Gln Leu Pro Tyr Ile
            740                 745                 750
Ala Leu Gly Met Leu Asn Phe Phe Leu Leu Cys Gly Leu Ser Val Phe
        755                 760                 765
Val Phe Gly Val Pro His Lys Gly Ser Phe Leu Thr Leu Thr Leu Ala
    770                 775                 780
Ala Leu Leu Tyr Ile Ile Ile Ala Thr Gly Met Gly Leu Leu Ile Ser
785                 790                 795                 800
Thr Phe Met Lys Ser Gln Ile Ala Ala Ile Phe Gly Thr Ala Ile Ile
                805                 810                 815
Thr Leu Ile Pro Ala Thr Gln Phe Ser Gly Met Ile Asp Pro Val Ala
            820                 825                 830
Ser Leu Glu Gly Pro Gly Arg Trp Ile Gly Glu Val Tyr Pro Thr Ser
        835                 840                 845
His Phe Leu Thr Ile Ala Arg Gly Thr Phe Ser Lys Ala Leu Asp Leu
    850                 855                 860
Thr Asp Leu Trp Gln Leu Phe Ile Pro Leu Leu Ile Ala Ile Pro Leu
865                 870                 875                 880
Val Met Gly Leu Ser Ile Leu Leu Leu Lys Lys Gln Glu Gly
                885                 890
```

<210> 435
<211> 355
<212> PRT
<213> Escherichia coli

<400> 435

```
Met Asp Lys Ser Lys Arg His Leu Ala Trp Trp Val Val Gly Leu Leu
 1               5                  10                  15
Ala Val Ala Ala Ile Val Ala Trp Trp Leu Leu Arg Pro Ala Gly Val
            20                  25                  30
Pro Glu Gly Phe Ala Val Ser Asn Gly Arg Ile Glu Ala Thr Glu Val
        35                  40                  45
Asp Ile Ala Ser Lys Ile Ala Gly Arg Ile Asp Thr Ile Leu Val Lys
    50                  55                  60
Glu Gly Lys Phe Val Arg Glu Gly Glu Val Leu Ala Lys Met Asp Thr
65                  70                  75                  80
Arg Val Leu Gln Glu Gln Arg Leu Glu Ala Ile Ala Gln Ile Lys Glu
                85                  90                  95
Ala Gln Ser Ala Val Ala Ala Ala Gln Ala Leu Leu Glu Gln Arg Gln
            100                 105                 110
```

Ser Glu Thr Arg Ala Ala Gln Ser Leu Val Asn Gln Arg Gln Ala Glu
        115   120   125
Leu Asp Ser Val Ala Lys Arg His Thr Arg Ser Arg Ser Leu Ala Gln
    130   135   140
Arg Gly Ala Ile Ser Ala Gln Gln Leu Asp Asp Asp Arg Ala Ala Ala
145   150   155   160
Glu Ser Ala Arg Ala Ala Leu Glu Ser Ala Lys Ala Gln Val Ser Ala
                165   170   175
Ser Lys Ala Ala Ile Glu Ala Ala Arg Thr Asn Ile Ile Gln Ala Gln
            180   185   190
Thr Arg Val Glu Ala Ala Gln Ala Thr Glu Arg Arg Ile Ala Ala Asp
        195   200   205
Ile Asp Asp Ser Glu Leu Lys Ala Pro Arg Asp Gly Arg Val Gln Tyr
    210   215   220
Arg Val Ala Glu Pro Gly Glu Val Leu Ala Ala Gly Gly Arg Val Leu
225   230   235   240
Asn Met Val Asp Leu Ser Asp Val Tyr Met Thr Phe Phe Leu Pro Thr
                245   250   255
Glu Gln Ala Gly Thr Leu Lys Leu Gly Gly Glu Ala Arg Leu Ile Leu
            260   265   270
Asp Ala Ala Pro Asp Leu Arg Ile Pro Ala Thr Ile Ser Phe Val Ala
        275   280   285
Ser Val Ala Gln Phe Thr Pro Lys Thr Val Glu Thr Ser Asp Glu Arg
    290   295   300
Leu Lys Leu Met Phe Arg Val Lys Ala Arg Ile Pro Pro Glu Leu Leu
305   310   315   320
Gln Gln His Leu Glu Tyr Val Lys Thr Gly Leu Pro Gly Val Ala Trp
                325   330   335
Val Arg Val Asn Glu Glu Leu Pro Trp Pro Asp Asp Leu Val Val Arg
            340   345   350
Leu Pro Gln
        355

<210> 436
<211> 235
<212> PRT
<213> Escherichia coli

<400> 436
Met Tyr Arg Tyr Leu Ser Ile Ala Ala Val Val Leu Ser Ala Ala Phe
1   5   10   15
Ser Gly Pro Ala Leu Ala Glu Gly Ile Asn Ser Phe Ser Gln Ala Lys
            20   25   30
Ala Ala Ala Val Lys Val His Ala Asp Ala Pro Gly Thr Phe Tyr Cys
        35   40   45
Gly Cys Lys Ile Asn Trp Gln Gly Lys Lys Gly Val Val Asp Leu Gln
    50   55   60
Ser Cys Gly Tyr Gln Val Arg Lys Asn Glu Asn Arg Ala Ser Arg Val
65   70   75   80
Glu Trp Glu His Val Val Pro Ala Trp Gln Phe Gly His Gln Arg Gln
                85   90   95
Cys Trp Gln Asp Gly Gly Arg Lys Asn Cys Ala Lys Asp Pro Val Tyr
            100   105   110
Arg Lys Met Glu Ser Asp Met His Asn Leu Gln Pro Ser Val Gly Glu
        115   120   125
Val Asn Gly Asp Arg Gly Asn Phe Met Tyr Ser Gln Trp Asn Gly Gly
    130   135   140
Glu Gly Gln Tyr Gly Gln Cys Ala Met Lys Val Asp Phe Lys Glu Lys
145   150   155   160
Ala Ala Glu Pro Pro Ala Arg Ala Arg Gly Ala Ile Ala Arg Thr Tyr

165 170 175
Phe Tyr Met Arg Asp Gln Tyr Asn Leu Thr Leu Ser Arg Gln Gln Thr
180 185 190
Gln Leu Phe Asn Ala Trp Asn Lys Met Tyr Pro Val Thr Asp Trp Glu
195 200 205
Cys Glu Arg Asp Glu Arg Ile Ala Lys Val Gln Gly Asn His Asn Pro
210 215 220
Tyr Val Gln Arg Ala Cys Gln Ala Arg Lys Ser
225 230 235

<210> 437
<211> 480
<212> PRT
<213> Escherichia coli

<400> 437
Met Ser Arg Arg Leu Arg Arg Thr Lys Ile Val Thr Thr Leu Gly Pro
1 5 10 15
Ala Thr Asp Arg Asp Asn Asn Leu Glu Lys Val Ile Ala Ala Gly Ala
20 25 30
Asn Val Val Arg Met Asn Phe Ser His Gly Ser Pro Glu Asp His Lys
35 40 45
Met Arg Ala Asp Lys Val Arg Glu Ile Ala Ala Lys Leu Gly Arg His
50 55 60
Val Ala Ile Leu Gly Asp Leu Gln Gly Pro Lys Ile Arg Val Ser Thr
65 70 75 80
Phe Lys Glu Gly Lys Val Phe Leu Asn Ile Gly Asp Lys Phe Leu Leu
85 90 95
Asp Ala Asn Leu Gly Lys Gly Glu Gly Asp Lys Glu Lys Val Gly Ile
100 105 110
Asp Tyr Lys Gly Leu Pro Ala Asp Val Val Pro Gly Asp Ile Leu Leu
115 120 125
Leu Asp Asp Gly Arg Val Gln Leu Lys Val Leu Glu Val Gln Gly Met
130 135 140
Lys Val Phe Thr Glu Val Thr Val Gly Gly Pro Leu Ser Asn Asn Lys
145 150 155 160
Gly Ile Asn Lys Leu Gly Gly Gly Leu Ser Ala Glu Ala Leu Thr Glu
165 170 175
Lys Asp Lys Ala Asp Ile Lys Thr Ala Ala Leu Ile Gly Val Asp Tyr
180 185 190
Leu Ala Val Ser Phe Pro Arg Cys Gly Glu Asp Leu Asn Tyr Ala Arg
195 200 205
Arg Leu Ala Arg Asp Ala Gly Cys Asp Ala Lys Ile Val Ala Lys Val
210 215 220
Glu Arg Ala Glu Ala Val Cys Ser Gln Asp Ala Met Asp Asp Ile Ile
225 230 235 240
Leu Ala Ser Asp Val Val Met Val Ala Arg Gly Asp Leu Gly Val Glu
245 250 255
Ile Gly Asp Pro Glu Leu Val Gly Ile Gln Lys Ala Leu Ile Arg Arg
260 265 270
Ala Arg Gln Leu Asn Arg Ala Val Ile Thr Ala Thr Gln Met Met Glu
275 280 285
Ser Met Ile Thr Asn Pro Met Pro Thr Arg Ala Glu Val Met Asp Val
290 295 300
Ala Asn Ala Val Leu Asp Gly Thr Asp Ala Val Met Leu Ser Ala Glu
305 310 315 320
Thr Ala Ala Gly Gln Tyr Pro Ser Glu Thr Val Ala Ala Met Ala Arg
325 330 335
Val Cys Leu Gly Ala Glu Lys Ile Pro Ser Ile Asn Val Ser Lys His
340 345 350

Arg Leu Asp Val Gln Phe Asp Asn Val Glu Glu Ala Ile Ala Met Ser
355 360 365
Ala Met Tyr Ala Ala Asn His Leu Lys Gly Val Thr Ala Ile Ile Thr
370 375 380
Met Thr Glu Ser Gly Arg Thr Ala Leu Met Thr Ser Arg Ile Ser Ser
385 390 395 400
Gly Leu Pro Ile Phe Ala Met Ser Arg His Glu Arg Thr Leu Asn Leu
405 410 415
Thr Ala Leu Tyr Arg Gly Val Thr Pro Val His Phe Asp Ser Ala Asn
420 425 430
Asp Gly Val Ala Ala Ala Ser Glu Ala Val Asn Leu Leu Arg Asp Lys
435 440 445
Gly Tyr Leu Met Ser Gly Asp Leu Val Ile Val Thr Gln Gly Asp Val
450 455 460
Met Ser Thr Val Gly Ser Thr Asn Thr Thr Arg Ile Leu Thr Val Glu
465 470 475 480

<210> 438
<211> 239
<212> PRT
<213> Escherichia coli

<400> 438
Met Ile Asn Val Leu Ile Ile Asp Asp Asp Ala Met Val Ala Glu Leu
1 5 10 15
Asn Arg Arg Tyr Val Ala Gln Ile Pro Gly Phe Gln Cys Cys Gly Thr
20 25 30
Ala Ser Thr Leu Glu Lys Ala Lys Glu Ile Ile Phe Asn Ser Asp Thr
35 40 45
Pro Ile Asp Leu Ile Leu Leu Asp Ile Tyr Met Gln Lys Glu Asn Gly
50 55 60
Leu Asp Leu Leu Pro Val Leu His Asn Ala Arg Cys Lys Ser Asp Val
65 70 75 80
Ile Val Ile Ser Ser Ala Ala Asp Ala Ala Thr Ile Lys Asp Ser Leu
85 90 95
His Tyr Gly Val Val Asp Tyr Leu Ile Lys Pro Phe Gln Ala Ser Arg
100 105 110
Phe Glu Glu Ala Leu Thr Gly Trp Arg Gln Lys Lys Met Ala Leu Glu
115 120 125
Lys His Gln Tyr Tyr Asp Gln Ala Glu Leu Asp Gln Leu Ile His Gly
130 135 140
Ser Ser Ser Asn Glu Gln Asp Pro Arg Arg Leu Pro Lys Gly Leu Thr
145 150 155 160
Pro Gln Thr Leu Arg Thr Leu Cys Gln Trp Ile Asp Ala His Gln Asp
165 170 175
Tyr Glu Phe Ser Thr Asp Glu Leu Ala Asn Glu Val Asn Ile Ser Arg
180 185 190
Val Ser Cys Arg Lys Tyr Leu Ile Trp Leu Val Asn Cys His Ile Leu
195 200 205
Phe Thr Ser Ile His Tyr Gly Val Thr Gly Arg Pro Val Tyr Arg Tyr
210 215 220
Arg Ile Gln Ala Glu His Tyr Ser Leu Leu Lys Gln Tyr Cys Gln
225 230 235

<210> 439
<211> 543
<212> PRT
<213> Escherichia coli

<400> 439
Met Arg His Ser Leu Pro Tyr Arg Met Leu Arg Lys Arg Pro Met Lys
1 5 10 15
Leu Ser Thr Thr Val Ile Leu Met Val Ser Ala Val Leu Phe Ser Val
20 25 30
Leu Leu Val Val His Leu Ile Tyr Phe Ser Gln Ile Ser Asp Met Thr
35 40 45
Arg Asp Gly Leu Ala Asn Lys Ala Leu Ala Val Ala Arg Thr Leu Ala
50 55 60
Asp Ser Pro Glu Ile Arg Gln Gly Leu Gln Lys Lys Pro Gln Glu Ser
65 70 75 80
Gly Ile Gln Ala Ile Ala Glu Ala Val Arg Lys Arg Asn Asp Leu Leu
85 90 95
Phe Ile Val Val Thr Asp Met Gln Ser Leu Arg Tyr Ser His Pro Glu
100 105 110
Ala Gln Arg Ile Gly Gln Pro Phe Lys Gly Asp Asp Ile Leu Lys Ala
115 120 125
Leu Asn Gly Glu Glu Asn Val Ala Ile Asn Arg Gly Phe Leu Ala Gln
130 135 140
Ala Leu Arg Val Phe Thr Pro Ile Tyr Asp Glu Asn His Lys Gln Ile
145 150 155 160
Gly Val Val Ala Ile Gly Leu Glu Leu Ser Arg Val Thr Gln Gln Ile
165 170 175
Asn Asp Ser Arg Trp Ser Ile Ile Trp Ser Val Leu Phe Gly Met Leu
180 185 190
Val Gly Leu Ile Gly Thr Cys Ile Leu Val Lys Val Leu Lys Lys Ile
195 200 205
Leu Phe Gly Leu Glu Pro Tyr Glu Ile Ser Thr Leu Phe Glu Gln Arg
210 215 220
Gln Ala Met Leu Gln Ser Ile Lys Glu Gly Val Val Ala Val Asp Asp
225 230 235 240
Arg Gly Glu Val Thr Leu Ile Asn Asp Ala Ala Gln Glu Leu Leu Asn
245 250 255
Tyr Arg Lys Ser Gln Asp Asp Glu Lys Leu Ser Thr Leu Ser His Ser
260 265 270
Trp Ser Gln Val Val Asp Val Ser Glu Val Leu Arg Asp Gly Thr Pro
275 280 285
Arg Arg Asp Glu Glu Ile Thr Ile Lys Asp Arg Leu Leu Leu Ile Asn
290 295 300
Thr Val Pro Val Arg Ser Asn Gly Val Ile Ile Gly Ala Ile Ser Thr
305 310 315 320
Phe Arg Asp Lys Thr Glu Val Arg Lys Leu Met Gln Arg Leu Asp Gly
325 330 335
Leu Val Asn Tyr Ala Asp Ala Leu Arg Glu Arg Ser His Glu Phe Met
340 345 350
Asn Lys Leu His Val Ile Leu Gly Leu Leu His Leu Lys Ser Tyr Lys
355 360 365
Gln Leu Glu Asp Tyr Ile Leu Lys Thr Ala Asn Asn Tyr Gln Glu Glu
370 375 380
Ile Gly Ser Leu Leu Gly Lys Ile Lys Ser Pro Val Ile Ala Gly Phe
385 390 395 400
Leu Ile Ser Lys Ile Asn Arg Ala Thr Asp Leu Gly His Thr Leu Ile
405 410 415
Leu Asn Ser Glu Ser Gln Leu Pro Asp Ser Gly Ser Glu Asp Gln Val
420 425 430
Ala Thr Leu Ile Thr Thr Leu Gly Asn Leu Ile Glu Asn Ala Leu Glu
435 440 445
Ala Leu Gly Pro Glu Pro Gly Gly Glu Ile Ser Val Thr Leu His Tyr
450 455 460
Arg His Gly Trp Leu His Cys Glu Val Asn Asp Asp Gly Pro Gly Ile
465 470 475 480

Ala Pro Asp Lys Ile Asp His Ile Phe Asp Lys Gly Val Ser Thr Lys
485 490 495
Gly Ser Glu Arg Gly Val Gly Leu Ala Leu Val Lys Gln Gln Val Glu
500 505 510
Asn Leu Gly Gly Ser Ile Ala Val Glu Ser Glu Pro Gly Ile Phe Thr
515 520 525
Gln Phe Phe Val Gln Ile Pro Trp Asp Gly Glu Arg Ser Asn Arg
530 535 540

<210> 440
<211> 328
<212> PRT
<213> Escherichia coli

<400> 440
Met Ser Val Pro Leu Ser Thr Trp Asn Leu Leu Arg Tyr Asn Asn Ser
1 5 10 15
Tyr Leu Gln Lys Val Thr Met Phe Pro Gln Cys Lys Phe Ser Arg Glu
20 25 30
Phe Leu His Pro Arg Tyr Trp Leu Thr Trp Phe Gly Leu Gly Val Leu
35 40 45
Trp Leu Trp Val Gln Leu Pro Tyr Pro Val Leu Cys Phe Leu Gly Thr
50 55 60
Arg Ile Gly Ala Met Ala Arg Pro Phe Leu Lys Arg Arg Glu Ser Ile
65 70 75 80
Ala Arg Lys Asn Leu Glu Leu Cys Phe Pro Gln His Ser Ala Glu Glu
85 90 95
Arg Glu Lys Met Ile Ala Glu Asn Phe Arg Ser Leu Gly Met Ala Leu
100 105 110
Val Glu Thr Gly Met Ala Trp Phe Trp Pro Asp Ser Arg Val Arg Lys
115 120 125
Trp Phe Asp Val Glu Gly Leu Asp Asn Leu Lys Arg Ala Gln Met Gln
130 135 140
Asn Arg Gly Val Met Val Val Gly Val His Phe Met Ser Leu Glu Leu
145 150 155 160
Gly Gly Arg Val Met Gly Leu Cys Gln Pro Met Met Ala Thr Tyr Arg
165 170 175
Pro His Asn Asn Gln Leu Met Glu Trp Val Gln Thr Arg Gly Arg Met
180 185 190
Arg Ser Asn Lys Ala Met Ile Gly Arg Asn Asn Leu Arg Gly Ile Val
195 200 205
Gly Ala Leu Lys Lys Gly Glu Ala Val Trp Phe Ala Pro Asp Gln Asp
210 215 220
Tyr Gly Arg Lys Gly Ser Ser Phe Ala Pro Phe Phe Ala Val Glu Asn
225 230 235 240
Val Ala Thr Thr Asn Gly Thr Tyr Val Leu Ser Arg Leu Ser Gly Ala
245 250 255
Ala Met Leu Thr Val Thr Met Val Arg Lys Ala Asp Tyr Ser Gly Tyr
260 265 270
Arg Leu Phe Ile Thr Pro Glu Met Glu Gly Tyr Pro Thr Asp Glu Asn
275 280 285
Gln Ala Ala Ala Tyr Met Asn Lys Ile Ile Glu Lys Glu Ile Met Arg
290 295 300
Ala Pro Glu Gln Tyr Leu Trp Ile His Arg Arg Phe Lys Thr Arg Pro
305 310 315 320
Val Gly Glu Ser Ser Leu Tyr Ile
325

<210> 441

<211> 87
<212> PRT
<213> Escherichia coli

<400> 441
Met Ala Asn Ile Lys Ser Ala Lys Lys Arg Ala Ile Gln Ser Glu Lys
1   5   10   15
Ala Arg Lys His Asn Ala Ser Arg Arg Ser Met Met Arg Thr Phe Ile
            20   25   30
Lys Lys Val Tyr Ala Ala Ile Glu Ala Gly Asp Lys Ala Ala Ala Gln
        35   40   45
Lys Ala Phe Asn Glu Met Gln Pro Ile Val Asp Arg Gln Ala Ala Lys
    50   55   60
Gly Leu Ile His Lys Asn Lys Ala Ala Arg His Lys Ala Asn Leu Thr
65   70   75   80
Ala Gln Ile Asn Lys Leu Ala
                85

<210> 442
<211> 430
<212> PRT
<213> Escherichia coli

<400> 442
Met Arg Tyr Asn Gly Leu Asn Asn Met Phe Phe Pro Leu Cys Leu Ile
1   5   10   15
Asn Asp Asn His Ser Val Thr Ser Pro Ser His Thr Lys Lys Thr Lys
            20   25   30
Ser Asp Asn Tyr Ser Lys His His Lys Asn Thr Leu Ile Asp Asn Lys
        35   40   45
Ala Leu Ser Leu Phe Lys Met Asp Asp His Glu Lys Val Ile Gly Leu
    50   55   60
Ile Gln Lys Met Lys Arg Ile Tyr Asp Ser Leu Pro Ser Gly Lys Ile
65   70   75   80
Thr Lys Glu Thr Asp Arg Lys Ile His Lys Tyr Phe Ile Asp Ile Ala
                85   90   95
Ser His Ala Asn Asn Lys Cys Asp Asp Arg Ile Thr Arg Arg Val Tyr
            100   105   110
Leu Asn Lys Asp Lys Glu Val Ser Ile Lys Val Val Tyr Phe Ile Asn
        115   120   125
Asn Val Thr Val His Asn Asn Thr Ile Glu Ile Pro Gln Thr Val Asn
    130   135   140
Gly Gly Tyr Asp Phe Ser His Leu Ser Leu Lys Gly Ile Val Ile Lys
145   150   155   160
Asp Glu Asp Leu Ser Asn Ser Asn Phe Ala Gly Cys Arg Leu Gln Asn
                165   170   175
Ala Ile Phe Gln Asp Cys Asn Met Tyr Lys Thr Asn Phe Asn Phe Ala
            180   185   190
Ile Met Glu Lys Ile Leu Phe Asp Asn Cys Ile Leu Asp Asp Ser Asn
        195   200   205
Phe Ala Gln Ile Lys Met Thr Asp Gly Thr Leu Asn Ser Cys Ser Ala
    210   215   220
Met His Val Gln Phe Tyr Asn Ala Thr Met Asn Arg Ala Asn Ile Lys
225   230   235   240
Asn Thr Phe Leu Asp Tyr Ser Asn Phe Tyr Met Ala Tyr Met Ala Glu
                245   250   255
Val Asn Leu Tyr Lys Val Ile Ala Pro Tyr Ile Asn Leu Phe Arg Ala
            260   265   270
Asp Leu Ser Phe Ser Lys Leu Asp Leu Ile Asn Phe Glu His Ala Asp
        275   280   285

Leu Ser Arg Val Asn Leu Asn Lys Ala Thr Leu Gln Asn Ile Asn Leu
    290   295   300
Ile Asp Ser Lys Leu Phe Phe Thr Arg Leu Thr Asn Thr Phe Leu Glu
305   310   315   320
Met Val Ile Cys Thr Asp Ser Asn Met Ala Asn Val Asn Phe Asn Asn
                325   330   335
Ala Asn Leu Ser Asn Cys His Phe Asn Cys Ser Val Leu Thr Lys Ala
            340   345   350
Trp Met Phe Asn Ile Arg Leu Tyr Arg Val Asn Phe Asp Glu Ala Ser
        355   360   365
Val Gln Gly Met Gly Ile Thr Ile Leu Arg Gly Glu Glu Asn Ile Ser
    370   375   380
Ile Asn Ser Asp Ile Leu Val Thr Leu Gln Lys Phe Phe Glu Glu Asp
385   390   395   400
Cys Ala Thr His Thr Gly Met Ser Gln Thr Glu Asp Asn Leu His Ala
                405   410   415
Val Ala Met Lys Ile Thr Ala Asp Ile Met Gln Asp Ala Asp
            420   425   430

<210> 443
<211> 883
<212> PRT
<213> Escherichia coli

<400> 443
Met Asn Glu Gln Tyr Ser Ala Leu Arg Ser Asn Val Ser Met Leu Gly
1   5   10   15
Lys Val Leu Gly Glu Thr Ile Lys Asp Ala Leu Gly Glu His Ile Leu
            20   25   30
Glu Arg Val Glu Thr Ile Arg Lys Leu Ser Lys Ser Ser Arg Ala Gly
        35   40   45
Asn Asp Ala Asn Arg Gln Glu Leu Leu Thr Thr Leu Gln Asn Leu Ser
    50   55   60
Asn Asp Glu Leu Leu Pro Val Ala Arg Ala Phe Ser Gln Phe Leu Asn
65   70   75   80
Leu Ala Asn Thr Ala Glu Gln Tyr His Ser Ile Ser Pro Lys Gly Glu
                85   90   95
Ala Ala Ser Asn Pro Glu Val Ile Ala Arg Thr Leu Arg Lys Leu Lys
            100   105   110
Asn Gln Pro Glu Leu Ser Glu Asp Thr Ile Lys Lys Ala Val Glu Ser
        115   120   125
Leu Ser Leu Glu Leu Val Leu Thr Ala His Pro Thr Glu Ile Thr Arg
    130   135   140
Arg Thr Leu Ile His Lys Met Val Glu Val Asn Ala Cys Leu Lys Gln
145   150   155   160
Leu Asp Asn Lys Asp Ile Ala Asp Tyr Glu His Asn Gln Leu Met Arg
                165   170   175
Arg Leu Arg Gln Leu Ile Ala Gln Ser Trp His Thr Asp Glu Ile Arg
            180   185   190
Lys Leu Arg Pro Ser Pro Val Asp Glu Ala Lys Trp Gly Phe Ala Val
        195   200   205
Val Glu Asn Ser Leu Trp Gln Gly Val Pro Asn Tyr Leu Arg Glu Leu
    210   215   220
Asn Glu Gln Leu Glu Glu Asn Leu Gly Tyr Lys Leu Pro Val Glu Phe
225   230   235   240
Val Pro Val Arg Phe Thr Ser Trp Met Gly Gly Asp Arg Asp Gly Asn
                245   250   255
Pro Asn Val Thr Ala Asp Ile Thr Arg His Val Leu Leu Leu Ser Arg
            260   265   270
Trp Lys Ala Thr Asp Leu Phe Leu Lys Asp Ile Gln Val Leu Val Ser

275 280 285
Glu Leu Ser Met Val Glu Ala Thr Pro Glu Leu Leu Ala Leu Val Gly
290 295 300
Glu Glu Gly Ala Ala Glu Pro Tyr Arg Tyr Leu Met Lys Asn Leu Arg
305 310 315 320
Ser Arg Leu Met Ala Thr Gln Ala Trp Leu Glu Ala Arg Leu Lys Gly
325 330 335
Glu Glu Leu Pro Lys Pro Glu Gly Leu Leu Thr Gln Asn Glu Glu Leu
340 345 350
Trp Glu Pro Leu Tyr Ala Cys Tyr Gln Ser Leu Gln Ala Cys Gly Met
355 360 365
Gly Ile Ile Ala Asn Gly Asp Leu Leu Asp Thr Leu Arg Arg Val Lys
370 375 380
Cys Phe Gly Val Pro Leu Val Arg Ile Asp Ile Arg Gln Glu Ser Thr
385 390 395 400
Arg His Thr Glu Ala Leu Gly Glu Leu Thr Arg Tyr Leu Gly Ile Gly
405 410 415
Asp Tyr Glu Ser Trp Ser Glu Ala Asp Lys Gln Ala Phe Leu Ile Arg
420 425 430
Glu Leu Asn Ser Lys Arg Pro Leu Leu Pro Arg Asn Trp Gln Pro Ser
435 440 445
Ala Glu Thr Arg Glu Val Leu Asp Thr Cys Gln Val Ile Ala Glu Ala
450 455 460
Pro Gln Gly Ser Ile Ala Ala Tyr Val Ile Ser Met Ala Lys Thr Pro
465 470 475 480
Ser Asp Val Leu Ala Val His Leu Leu Leu Lys Glu Ala Gly Ile Gly
485 490 495
Phe Ala Met Pro Val Ala Pro Leu Phe Glu Thr Leu Asp Asp Leu Asn
500 505 510
Asn Ala Asn Asp Val Met Thr Gln Leu Leu Asn Ile Asp Trp Tyr Arg
515 520 525
Gly Leu Ile Gln Gly Lys Gln Met Val Met Ile Gly Tyr Ser Asp Ser
530 535 540
Ala Lys Asp Ala Gly Val Met Ala Ala Ser Trp Ala Gln Tyr Gln Ala
545 550 555 560
Gln Asp Ala Leu Ile Lys Thr Cys Glu Lys Ala Gly Ile Glu Leu Thr
565 570 575
Leu Phe His Gly Arg Gly Gly Ser Ile Gly Arg Gly Gly Ala Pro Ala
580 585 590
His Ala Ala Leu Leu Ser Gln Pro Pro Gly Ser Leu Lys Gly Gly Leu
595 600 605
Arg Val Thr Glu Gln Gly Glu Met Ile Arg Phe Lys Tyr Gly Leu Pro
610 615 620
Glu Ile Thr Val Ser Ser Leu Ser Leu Tyr Thr Gly Ala Ile Leu Glu
625 630 635 640
Ala Asn Leu Leu Pro Pro Pro Glu Pro Lys Glu Ser Trp Arg Arg Ile
645 650 655
Met Asp Glu Leu Ser Val Ile Ser Cys Asp Val Tyr Arg Gly Tyr Val
660 665 670
Arg Glu Asn Lys Asp Phe Val Pro Tyr Phe Arg Ser Ala Thr Pro Glu
675 680 685
Gln Glu Leu Gly Lys Leu Pro Leu Gly Ser Arg Pro Ala Lys Arg Arg
690 695 700
Pro Thr Gly Gly Val Glu Ser Leu Arg Ala Ile Pro Trp Ile Phe Ala
705 710 715 720
Trp Thr Gln Asn Arg Leu Met Leu Pro Ala Trp Leu Gly Ala Gly Thr
725 730 735
Ala Leu Gln Lys Val Val Glu Asp Gly Lys Gln Ser Glu Leu Glu Ala
740 745 750
Met Cys Arg Asp Trp Pro Phe Phe Ser Thr Arg Leu Gly Met Leu Glu
755 760 765

Met Val Phe Ala Lys Ala Asp Leu Trp Leu Ala Glu Tyr Tyr Asp Gln
770 775 780
Arg Leu Val Asp Lys Ala Leu Trp Pro Leu Gly Lys Glu Leu Arg Asn
785 790 795 800
Leu Gln Glu Glu Asp Ile Lys Val Val Leu Ala Ile Ala Asn Asp Ser
805 810 815
His Leu Met Ala Asp Leu Pro Trp Ile Ala Glu Ser Ile Gln Leu Arg
820 825 830
Asn Ile Tyr Thr Asp Pro Leu Asn Val Leu Gln Ala Glu Leu Leu His
835 840 845
Arg Ser Arg Gln Ala Glu Lys Glu Gly Gln Glu Pro Asp Pro Arg Val
850 855 860
Glu Gln Ala Leu Met Val Thr Ile Ala Gly Ile Ala Ala Gly Met Arg
865 870 875 880
Asn Thr Gly

<210> 444
<211> 663
<212> PRT
<213> Escherichia coli

<400> 444
Met Ser Ser Arg Lys Glu Leu Ala Asn Ala Ile Arg Ala Leu Ser Met
1 5 10 15
Asp Ala Val Gln Lys Ala Lys Ser Gly His Pro Gly Ala Pro Met Gly
20 25 30
Met Ala Asp Ile Ala Glu Val Leu Trp Arg Asp Phe Leu Lys His Asn
35 40 45
Pro Gln Asn Pro Ser Trp Ala Asp Arg Asp Arg Phe Val Leu Ser Asn
50 55 60
Gly His Gly Ser Met Leu Ile Tyr Ser Leu Leu His Leu Thr Gly Tyr
65 70 75 80
Asp Leu Pro Met Glu Glu Leu Lys Asn Phe Arg Gln Leu His Ser Lys
85 90 95
Thr Pro Gly His Pro Glu Val Gly Tyr Thr Ala Gly Val Glu Thr Thr
100 105 110
Thr Gly Pro Leu Gly Gln Gly Ile Ala Asn Ala Val Gly Met Ala Ile
115 120 125
Ala Glu Lys Thr Leu Ala Ala Gln Phe Asn Arg Pro Gly His Asp Ile
130 135 140
Val Asp His Tyr Thr Tyr Ala Phe Met Gly Asp Gly Cys Met Met Glu
145 150 155 160
Gly Ile Ser His Glu Val Cys Ser Leu Ala Gly Thr Leu Lys Leu Gly
165 170 175
Lys Leu Ile Ala Phe Tyr Asp Asp Asn Gly Ile Ser Ile Asp Gly His
180 185 190
Val Glu Gly Trp Phe Thr Asp Asp Thr Ala Met Arg Phe Glu Ala Tyr
195 200 205
Gly Trp His Val Ile Arg Asp Ile Asp Gly His Asp Ala Ala Ser Ile
210 215 220
Lys Arg Ala Val Glu Glu Ala Arg Ala Val Thr Asp Lys Pro Ser Leu
225 230 235 240
Leu Met Cys Lys Thr Ile Ile Gly Phe Gly Ser Pro Asn Lys Ala Gly
245 250 255
Thr His Asp Ser His Gly Ala Pro Leu Gly Asp Ala Glu Ile Ala Leu
260 265 270
Thr Arg Glu Gln Leu Gly Trp Lys Tyr Ala Pro Phe Glu Ile Pro Ser
275 280 285
Glu Ile Tyr Ala Gln Trp Asp Ala Lys Glu Ala Gly Gln Ala Lys Glu

290 295 300
Ser Ala Trp Asn Glu Lys Phe Ala Ala Tyr Ala Lys Ala Tyr Pro Gln
305 310 315 320
Glu Ala Ala Glu Phe Thr Arg Arg Met Lys Gly Glu Met Pro Ser Asp
325 330 335
Phe Asp Ala Lys Ala Lys Glu Phe Ile Ala Lys Leu Gln Ala Asn Pro
340 345 350
Ala Lys Ile Ala Ser Arg Lys Ala Ser Gln Asn Ala Ile Glu Ala Phe
355 360 365
Gly Pro Leu Leu Pro Glu Phe Leu Gly Gly Ser Ala Asp Leu Ala Pro
370 375 380
Ser Asn Leu Thr Leu Trp Ser Gly Ser Lys Ala Ile Asn Glu Asp Ala
385 390 395 400
Ala Gly Asn Tyr Ile His Tyr Gly Val Arg Glu Phe Gly Met Thr Ala
405 410 415
Ile Ala Asn Gly Ile Ser Leu His Gly Gly Phe Leu Pro Tyr Thr Ser
420 425 430
Thr Phe Leu Met Phe Val Glu Tyr Ala Arg Asn Ala Val Arg Met Ala
435 440 445
Ala Leu Met Lys Gln Arg Gln Val Met Val Tyr Thr His Asp Ser Ile
450 455 460
Gly Leu Gly Glu Asp Gly Pro Thr His Gln Pro Val Glu Gln Val Ala
465 470 475 480
Ser Leu Arg Val Thr Pro Asn Met Ser Thr Trp Arg Pro Cys Asp Gln
485 490 495
Val Glu Ser Ala Val Ala Trp Lys Tyr Gly Val Glu Arg Gln Asp Gly
500 505 510
Pro Thr Ala Leu Ile Leu Ser Arg Gln Asn Leu Ala Gln Gln Glu Arg
515 520 525
Thr Glu Glu Gln Leu Ala Asn Ile Ala Arg Gly Gly Tyr Val Leu Lys
530 535 540
Asp Cys Ala Gly Gln Pro Glu Leu Ile Phe Ile Ala Thr Gly Ser Glu
545 550 555 560
Val Glu Leu Ala Val Ala Ala Tyr Glu Lys Leu Thr Ala Glu Gly Val
565 570 575
Lys Ala Arg Val Val Ser Met Ser Ser Thr Asp Ala Phe Asp Lys Gln
580 585 590
Asp Ala Ala Tyr Arg Glu Ser Val Leu Pro Lys Ala Val Thr Ala Arg
595 600 605
Val Ala Val Glu Ala Gly Ile Ala Asp Tyr Trp Tyr Lys Tyr Val Gly
610 615 620
Leu Asn Gly Ala Ile Val Gly Met Thr Thr Phe Gly Glu Ser Ala Pro
625 630 635 640
Ala Glu Leu Leu Phe Glu Glu Phe Gly Phe Thr Val Asp Asn Val Val
645 650 655
Ala Lys Ala Lys Glu Leu Leu
660

<210> 445
<211> 152
<212> PRT
<213> Escherichia coli

<400> 445
Met Phe Arg Gly Ala Thr Leu Val Asn Leu Asp Ser Lys Gly Arg Leu
1 5 10 15
Ser Val Pro Thr Arg Tyr Arg Glu Gln Leu Leu Glu Asn Ala Ala Gly
20 25 30
Gln Met Val Cys Thr Ile Asp Ile Tyr His Pro Cys Leu Leu Leu Tyr
35 40 45

Pro Leu Pro Glu Trp Glu Ile Ile Glu Gln Lys Leu Ser Arg Leu Ser
50 55 60
Ser Met Asn Pro Val Glu Arg Arg Val Gln Arg Leu Leu Leu Gly His
65 70 75 80
Ala Ser Glu Cys Gln Met Asp Gly Ala Gly Arg Leu Leu Ile Ala Pro
85 90 95
Val Leu Arg Gln His Ala Gly Leu Thr Lys Glu Val Met Leu Val Gly
100 105 110
Gln Phe Asn Lys Phe Glu Leu Trp Asp Glu Thr Thr Trp His Gln Gln
115 120 125
Val Lys Glu Asp Ile Asp Ala Glu Gln Leu Ala Thr Gly Asp Leu Ser
130 135 140
Glu Arg Leu Gln Asp Leu Ser Leu
145 150

<210> 446
<211> 313
<212> PRT
<213> Escherichia coli

<400> 446
Met Met Glu Asn Tyr Lys His Thr Thr Val Leu Leu Asp Glu Ala Val
1 5 10 15
Asn Gly Leu Asn Ile Arg Pro Asp Gly Ile Tyr Ile Asp Gly Thr Phe
20 25 30
Gly Arg Gly Gly His Ser Arg Leu Ile Leu Ser Gln Leu Gly Glu Glu
35 40 45
Gly Arg Leu Leu Ala Ile Asp Arg Asp Pro Gln Ala Ile Ala Val Ala
50 55 60
Lys Thr Ile Asp Asp Pro Arg Phe Ser Ile Ile His Gly Pro Phe Ser
65 70 75 80
Ala Leu Gly Glu Tyr Val Ala Glu Arg Asp Leu Ile Gly Lys Ile Asp
85 90 95
Gly Ile Leu Leu Asp Leu Gly Val Ser Ser Pro Gln Leu Asp Asp Ala
100 105 110
Glu Arg Gly Phe Ser Phe Met Arg Asp Gly Pro Leu Asp Met Arg Met
115 120 125
Asp Pro Thr Arg Gly Gln Ser Ala Ala Glu Trp Leu Gln Thr Ala Glu
130 135 140
Glu Ala Asp Ile Ala Trp Val Leu Lys Thr Tyr Gly Glu Glu Arg Phe
145 150 155 160
Ala Lys Arg Ile Ala Arg Ala Ile Val Glu Arg Asn Arg Glu Gln Pro
165 170 175
Met Thr Arg Thr Lys Glu Leu Ala Glu Val Val Ala Ala Ala Thr Pro
180 185 190
Val Lys Asp Lys Phe Lys His Pro Ala Thr Arg Thr Phe Gln Ala Val
195 200 205
Arg Ile Trp Val Asn Ser Glu Leu Glu Glu Ile Glu Gln Ala Leu Lys
210 215 220
Ser Ser Leu Asn Val Leu Ala Pro Gly Gly Arg Leu Ser Ile Ile Ser
225 230 235 240
Phe His Ser Leu Glu Asp Arg Ile Val Lys Arg Phe Met Arg Glu Asn
245 250 255
Ser Arg Gly Pro Gln Val Pro Ala Gly Leu Pro Met Thr Glu Glu Gln
260 265 270
Leu Lys Lys Leu Gly Gly Arg Gln Leu Arg Ala Leu Gly Lys Leu Met
275 280 285
Pro Gly Glu Glu Glu Val Ala Glu Asn Pro Arg Ala Arg Ser Ser Val
290 295 300
Leu Arg Ile Ala Glu Arg Thr Asn Ala

305 310

<210> 447
<211> 121
<212> PRT
<213> Escherichia coli

<400> 447
Met Ile Ser Arg Val Thr Glu Ala Leu Ser Lys Val Lys Gly Ser Met
1 5 10 15
Gly Ser His Glu Arg His Ala Leu Pro Gly Val Ile Gly Asp Asp Leu
20 25 30
Leu Arg Phe Gly Lys Leu Pro Leu Cys Leu Phe Ile Cys Ile Ile Leu
35 40 45
Thr Ala Val Thr Val Val Thr Thr Ala His His Thr Arg Leu Leu Thr
50 55 60
Ala Gln Arg Glu Gln Leu Val Leu Glu Arg Asp Ala Leu Asp Ile Glu
65 70 75 80
Trp Arg Asn Leu Ile Leu Glu Glu Asn Ala Leu Gly Asp His Ser Arg
85 90 95
Val Glu Arg Ile Ala Thr Glu Lys Leu Gln Met Gln His Val Asp Pro
100 105 110
Ser Gln Glu Asn Ile Val Val Gln Lys
115 120

<210> 448
<211> 588
<212> PRT
<213> Escherichia coli

<400> 448
Met Lys Ala Ala Ala Lys Thr Gln Lys Pro Lys Arg Gln Glu Glu His
1 5 10 15
Ala Asn Phe Ile Ser Trp Arg Phe Ala Leu Leu Cys Gly Cys Ile Leu
20 25 30
Leu Ala Leu Ala Phe Leu Leu Gly Arg Val Ala Trp Leu Gln Val Ile
35 40 45
Ser Pro Asp Met Leu Val Lys Glu Gly Asp Met Arg Ser Leu Arg Val
50 55 60
Gln Gln Val Ser Thr Ser Arg Gly Met Ile Thr Asp Arg Ser Gly Arg
65 70 75 80
Pro Leu Ala Val Ser Val Pro Val Lys Ala Ile Trp Ala Asp Pro Lys
85 90 95
Glu Val His Asp Ala Gly Gly Ile Ser Val Gly Asp Arg Trp Lys Ala
100 105 110
Leu Ala Asn Ala Leu Asn Ile Pro Leu Asp Gln Leu Ser Ala Arg Ile
115 120 125
Asn Ala Asn Pro Lys Gly Arg Phe Ile Tyr Leu Ala Arg Gln Val Asn
130 135 140
Pro Asp Met Ala Asp Tyr Ile Lys Lys Leu Lys Leu Pro Gly Ile His
145 150 155 160
Leu Arg Glu Glu Ser Arg Arg Tyr Tyr Pro Ser Gly Glu Val Thr Ala
165 170 175
His Leu Ile Gly Phe Thr Asn Val Asp Ser Gln Gly Ile Glu Gly Val
180 185 190
Glu Lys Ser Phe Asp Lys Trp Leu Thr Gly Gln Pro Gly Glu Arg Ile
195 200 205
Val Arg Lys Asp Arg Tyr Gly Arg Val Ile Glu Asp Ile Ser Ser Thr
210 215 220

Asp Ser Gln Ala Ala His Asn Leu Ala Leu Ser Ile Asp Glu Arg Leu
225 230 235 240
Gln Ala Leu Val Tyr Arg Glu Leu Asn Asn Ala Val Ala Phe Asn Lys
245 250 255
Ala Glu Ser Gly Ser Ala Val Leu Val Asp Val Asn Thr Gly Glu Val
260 265 270
Leu Ala Met Ala Asn Ser Pro Ser Tyr Asn Pro Asn Asn Leu Ser Gly
275 280 285
Thr Pro Lys Glu Ala Met Arg Asn Arg Thr Ile Thr Asp Val Phe Glu
290 295 300
Pro Gly Ser Thr Val Lys Pro Met Val Val Met Thr Ala Leu Gln Arg
305 310 315 320
Gly Val Val Arg Glu Asn Ser Val Leu Asn Thr Ile Pro Tyr Arg Ile
325 330 335
Asn Gly His Glu Ile Lys Asp Val Ala Arg Tyr Ser Glu Leu Thr Leu
340 345 350
Thr Gly Val Leu Gln Lys Ser Ser Asn Val Gly Val Ser Lys Leu Ala
355 360 365
Leu Ala Met Pro Ser Ser Ala Leu Val Asp Thr Tyr Ser Arg Phe Gly
370 375 380
Leu Gly Lys Ala Thr Asn Leu Gly Leu Val Gly Glu Arg Ser Gly Leu
385 390 395 400
Tyr Pro Gln Lys Gln Arg Trp Ser Asp Ile Glu Arg Ala Thr Phe Ser
405 410 415
Phe Gly Tyr Gly Leu Met Val Thr Pro Leu Gln Leu Ala Arg Val Tyr
420 425 430
Ala Thr Ile Gly Ser Tyr Gly Ile Tyr Arg Pro Leu Ser Ile Thr Lys
435 440 445
Val Asp Pro Pro Val Pro Gly Glu Arg Val Phe Pro Glu Ser Ile Val
450 455 460
Arg Thr Val Val His Met Met Glu Ser Val Ala Leu Pro Gly Gly Gly
465 470 475 480
Gly Val Lys Ala Ala Ile Lys Gly Tyr Arg Ile Ala Ile Lys Thr Gly
485 490 495
Thr Ala Lys Lys Val Gly Pro Asp Gly Arg Tyr Ile Asn Lys Tyr Ile
500 505 510
Ala Tyr Thr Ala Gly Val Ala Pro Ala Ser Gln Pro Arg Phe Ala Leu
515 520 525
Val Val Val Ile Asn Asp Pro Gln Ala Gly Lys Tyr Tyr Gly Gly Ala
530 535 540
Val Ser Ala Pro Val Phe Gly Ala Ile Met Gly Gly Val Leu Arg Thr
545 550 555 560
Met Asn Ile Glu Pro Asp Ala Leu Thr Thr Gly Asp Lys Asn Glu Phe
565 570 575
Val Ile Asn Gln Gly Glu Gly Thr Gly Gly Arg Ser
580 585

<210> 449
<211> 495
<212> PRT
<213> Escherichia coli

<400> 449
Met Ala Asp Arg Asn Leu Arg Asp Leu Leu Ala Pro Trp Val Pro Asp
1 5 10 15
Ala Pro Ser Arg Ala Leu Arg Glu Met Thr Leu Asp Ser Arg Val Ala
20 25 30
Ala Ala Gly Asp Leu Phe Val Ala Val Val Gly His Gln Ala Asp Gly
35 40 45
Arg Arg Tyr Ile Pro Gln Ala Ile Ala Gln Gly Val Ala Ala Ile Ile

```
        50                  55                  60
Ala Glu Ala Lys Asp Glu Ala Thr Asp Gly Glu Ile Arg Glu Met His
65                  70                  75                  80
Gly Val Pro Val Ile Tyr Leu Ser Gln Leu Asn Glu Arg Leu Ser Ala
                85                  90                  95
Leu Ala Gly Arg Phe Tyr His Glu Pro Ser Asp Asn Leu Arg Leu Val
            100                 105                 110
Gly Val Thr Gly Thr Asn Gly Lys Thr Thr Thr Thr Gln Leu Leu Ala
        115                 120                 125
Gln Trp Ser Gln Leu Leu Gly Glu Ile Ser Ala Val Met Gly Thr Val
    130                 135                 140
Gly Asn Gly Leu Leu Gly Lys Val Ile Pro Thr Glu Asn Thr Thr Gly
145                 150                 155                 160
Ser Ala Val Asp Val Gln His Glu Leu Ala Gly Leu Val Asp Gln Gly
                165                 170                 175
Ala Thr Phe Cys Ala Met Glu Val Ser Ser His Gly Leu Val Gln His
            180                 185                 190
Arg Val Ala Ala Leu Lys Phe Ala Ala Ser Val Phe Thr Asn Leu Ser
        195                 200                 205
Arg Asp His Leu Asp Tyr His Gly Asp Met Glu His Tyr Glu Ala Ala
    210                 215                 220
Lys Trp Leu Leu Tyr Ser Glu His His Cys Gly Gln Ala Ile Ile Asn
225                 230                 235                 240
Ala Asp Asp Glu Val Gly Arg Arg Trp Leu Ala Lys Leu Pro Asp Ala
                245                 250                 255
Val Ala Val Ser Met Glu Asp His Ile Asn Pro Asn Cys His Gly Arg
            260                 265                 270
Trp Leu Lys Ala Thr Glu Val Asn Tyr His Asp Ser Gly Ala Thr Ile
        275                 280                 285
Arg Phe Ser Ser Ser Trp Gly Asp Gly Glu Ile Glu Ser His Leu Met
    290                 295                 300
Gly Ala Phe Asn Val Ser Asn Leu Leu Leu Ala Leu Ala Thr Leu Leu
305                 310                 315                 320
Ala Leu Gly Tyr Pro Leu Ala Asp Leu Leu Lys Thr Ala Ala Arg Leu
                325                 330                 335
Gln Pro Val Cys Gly Arg Met Glu Val Phe Thr Ala Pro Gly Lys Pro
            340                 345                 350
Thr Val Val Val Asp Tyr Ala His Thr Pro Asp Ala Leu Glu Lys Ala
        355                 360                 365
Leu Gln Ala Ala Arg Leu His Cys Ala Gly Lys Leu Trp Cys Val Phe
    370                 375                 380
Gly Cys Gly Gly Asp Arg Asp Lys Gly Lys Arg Pro Leu Met Gly Ala
385                 390                 395                 400
Ile Ala Glu Glu Phe Ala Asp Val Ala Val Val Thr Asp Asp Asn Pro
                405                 410                 415
Arg Thr Glu Glu Pro Arg Ala Ile Ile Asn Asp Ile Leu Ala Gly Met
            420                 425                 430
Leu Asp Ala Gly His Ala Lys Val Met Glu Gly Arg Ala Glu Ala Val
        435                 440                 445
Thr Cys Ala Val Met Gln Ala Lys Glu Asn Asp Val Val Leu Val Ala
    450                 455                 460
Gly Lys Gly His Glu Asp Tyr Gln Ile Val Gly Asn Gln Arg Leu Asp
465                 470                 475                 480
Tyr Ser Asp Arg Val Thr Val Ala Arg Leu Leu Gly Val Ile Ala
                485                 490                 495
```

<210> 450
<211> 452
<212> PRT
<213> Escherichia coli

<400> 450
Met Ile Ser Val Thr Leu Ser Gln Leu Thr Asp Ile Leu Asn Gly Glu
1 5 10 15
Leu Gln Gly Ala Asp Ile Thr Leu Asp Ala Val Thr Thr Asp Thr Arg
20 25 30
Lys Leu Thr Pro Gly Cys Leu Phe Val Ala Leu Lys Gly Glu Arg Phe
35 40 45
Asp Ala His Asp Phe Ala Asp Gln Ala Lys Ala Gly Gly Ala Gly Ala
50 55 60
Leu Leu Val Ser Arg Pro Leu Asp Ile Asp Leu Pro Gln Leu Ile Val
65 70 75 80
Lys Asp Thr Arg Leu Ala Phe Gly Glu Leu Ala Ala Trp Val Arg Gln
85 90 95
Gln Val Pro Ala Arg Val Val Ala Leu Thr Gly Ser Ser Gly Lys Thr
100 105 110
Ser Val Lys Glu Met Thr Ala Ala Ile Leu Ser Gln Cys Gly Asn Thr
115 120 125
Leu Tyr Thr Ala Gly Asn Leu Asn Asn Asp Ile Gly Val Pro Met Thr
130 135 140
Leu Leu Arg Leu Thr Pro Glu Tyr Asp Tyr Ala Val Ile Glu Leu Gly
145 150 155 160
Ala Asn His Gln Gly Glu Ile Ala Trp Thr Val Ser Leu Thr Arg Pro
165 170 175
Glu Ala Ala Leu Val Asn Asn Leu Ala Ala Ala His Leu Glu Gly Phe
180 185 190
Gly Ser Leu Ala Gly Val Ala Lys Ala Lys Gly Glu Ile Phe Ser Gly
195 200 205
Leu Pro Glu Asn Gly Ile Ala Ile Met Asn Ala Asp Asn Asn Asp Trp
210 215 220
Leu Asn Trp Gln Ser Val Ile Gly Ser Arg Lys Val Trp Arg Phe Ser
225 230 235 240
Pro Asn Ala Ala Asn Ser Asp Phe Thr Ala Thr Asn Ile His Val Thr
245 250 255
Ser His Gly Thr Glu Phe Thr Leu Gln Thr Pro Thr Gly Ser Val Asp
260 265 270
Val Leu Leu Pro Leu Pro Gly Arg His Asn Ile Ala Asn Ala Leu Ala
275 280 285
Ala Ala Ala Leu Ser Met Ser Val Gly Ala Thr Leu Asp Ala Ile Lys
290 295 300
Ala Gly Leu Ala Asn Leu Lys Ala Val Pro Gly Arg Leu Phe Pro Ile
305 310 315 320
Gln Leu Ala Glu Asn Gln Leu Leu Leu Asp Asp Ser Tyr Asn Ala Asn
325 330 335
Val Gly Ser Met Thr Ala Ala Val Gln Val Leu Ala Glu Met Pro Gly
340 345 350
Tyr Arg Val Leu Val Val Gly Asp Met Ala Glu Leu Gly Ala Glu Ser
355 360 365
Glu Ala Cys His Val Gln Val Gly Glu Ala Ala Lys Ala Ala Gly Ile
370 375 380
Asp Arg Val Leu Ser Val Gly Lys Gln Ser His Ala Ile Ser Thr Ala
385 390 395 400
Ser Gly Val Gly Glu His Phe Ala Asp Lys Thr Ala Leu Ile Thr Arg
405 410 415
Leu Lys Leu Leu Ile Ala Glu Gln Gln Val Ile Thr Ile Leu Val Lys
420 425 430
Gly Ser Arg Ser Ala Ala Met Glu Glu Val Val Arg Ala Leu Gln Glu
435 440 445
Asn Gly Thr Cys
450

<210> 451
<211> 360
<212> PRT
<213> Escherichia coli

<400> 451
Met Leu Val Trp Leu Ala Glu His Leu Val Lys Tyr Tyr Ser Gly Phe
1 5 10 15
Asn Val Phe Ser Tyr Leu Thr Phe Arg Ala Ile Val Ser Leu Leu Thr
20 25 30
Ala Leu Phe Ile Ser Leu Trp Met Gly Pro Arg Met Ile Ala His Leu
35 40 45
Gln Lys Leu Ser Phe Gly Gln Val Val Arg Asn Asp Gly Pro Glu Ser
50 55 60
His Phe Ser Lys Arg Gly Thr Pro Thr Met Gly Gly Ile Met Ile Leu
65 70 75 80
Thr Ala Ile Val Ile Ser Val Leu Leu Trp Ala Tyr Pro Ser Asn Pro
85 90 95
Tyr Val Trp Cys Val Leu Val Val Leu Val Gly Tyr Gly Val Ile Gly
100 105 110
Phe Val Asp Asp Tyr Arg Lys Val Val Arg Lys Asp Thr Lys Gly Leu
115 120 125
Ile Ala Arg Trp Lys Tyr Phe Trp Met Ser Val Ile Ala Leu Gly Val
130 135 140
Ala Phe Ala Leu Tyr Leu Ala Gly Lys Asp Thr Pro Ala Thr Gln Leu
145 150 155 160
Val Val Pro Phe Phe Lys Asp Val Met Pro Gln Leu Gly Leu Phe Tyr
165 170 175
Ile Leu Leu Ala Tyr Phe Val Ile Val Gly Thr Gly Asn Ala Val Asn
180 185 190
Leu Thr Asp Gly Leu Asp Gly Leu Ala Ile Met Pro Thr Val Phe Val
195 200 205
Ala Gly Gly Phe Ala Leu Val Ala Trp Ala Thr Gly Asn Met Asn Phe
210 215 220
Ala Ser Tyr Leu His Ile Pro Tyr Leu Arg His Ala Gly Glu Leu Val
225 230 235 240
Ile Val Cys Thr Ala Ile Val Gly Ala Gly Leu Gly Phe Leu Trp Phe
245 250 255
Asn Thr Tyr Pro Ala Gln Val Phe Met Gly Asp Val Gly Ser Leu Ala
260 265 270
Leu Gly Gly Ala Leu Gly Ile Ile Ala Val Leu Leu Arg Gln Glu Phe
275 280 285
Leu Leu Val Ile Met Gly Gly Val Phe Val Val Glu Thr Leu Ser Val
290 295 300
Ile Leu Gln Val Gly Ser Phe Lys Leu Arg Gly Gln Arg Ile Phe Arg
305 310 315 320
Met Ala Pro Ile His His His Tyr Glu Leu Lys Gly Trp Pro Glu Pro
325 330 335
Arg Val Ile Val Arg Phe Trp Ile Ile Ser Leu Met Leu Val Leu Ile
340 345 350
Gly Leu Ala Thr Leu Lys Val Arg
355 360

<210> 452
<211> 438
<212> PRT
<213> Escherichia coli

<400> 452

Met Ala Asp Tyr Gln Gly Lys Asn Val Val Ile Ile Gly Leu Gly Leu
1 5 10 15
Thr Gly Leu Ser Cys Val Asp Phe Phe Leu Ala Arg Gly Val Thr Pro
20 25 30
Arg Val Met Asp Thr Arg Met Thr Pro Pro Gly Leu Asp Lys Leu Pro
35 40 45
Glu Ala Val Glu Arg His Thr Gly Ser Leu Asn Asp Glu Trp Leu Met
50 55 60
Ala Ala Asp Leu Ile Val Ala Ser Pro Gly Ile Ala Leu Ala His Pro
65 70 75 80
Ser Leu Ser Ala Ala Ala Asp Ala Gly Ile Glu Ile Val Gly Asp Ile
85 90 95
Glu Leu Phe Cys Arg Glu Ala Gln Ala Pro Ile Val Ala Ile Thr Gly
100 105 110
Ser Asn Gly Lys Ser Thr Val Thr Thr Leu Val Gly Glu Met Ala Lys
115 120 125
Ala Ala Gly Val Asn Val Gly Val Gly Gly Asn Ile Gly Leu Pro Ala
130 135 140
Leu Met Leu Leu Asp Asp Glu Cys Glu Leu Tyr Val Leu Glu Leu Ser
145 150 155 160
Ser Phe Gln Leu Glu Thr Thr Ser Ser Leu Gln Ala Val Ala Ala Thr
165 170 175
Ile Leu Asn Val Thr Glu Asp His Met Asp Arg Tyr Pro Phe Gly Leu
180 185 190
Gln Gln Tyr Arg Ala Ala Lys Leu Arg Ile Tyr Glu Asn Ala Lys Val
195 200 205
Cys Val Val Asn Ala Asp Asp Ala Leu Thr Met Pro Ile Arg Gly Ala
210 215 220
Asp Glu Arg Cys Val Ser Phe Gly Val Asn Met Gly Asp Tyr His Leu
225 230 235 240
Asn His Gln Gln Gly Glu Thr Trp Leu Arg Val Lys Gly Glu Lys Val
245 250 255
Leu Asn Val Lys Glu Met Lys Leu Ser Gly Gln His Asn Tyr Thr Asn
260 265 270
Ala Leu Ala Ala Leu Ala Leu Ala Asp Ala Ala Gly Leu Pro Arg Ala
275 280 285
Ser Ser Leu Lys Ala Leu Thr Thr Phe Thr Gly Leu Pro His Arg Phe
290 295 300
Glu Val Val Leu Glu His Asn Gly Val Arg Trp Ile Asn Asp Ser Lys
305 310 315 320
Ala Thr Asn Val Gly Ser Thr Glu Ala Ala Leu Asn Gly Leu His Val
325 330 335
Asp Gly Thr Leu His Leu Leu Leu Gly Gly Asp Gly Lys Ser Ala Asp
340 345 350
Phe Ser Pro Leu Ala Arg Tyr Leu Asn Gly Asp Asn Val Arg Leu Tyr
355 360 365
Cys Phe Gly Arg Asp Gly Ala Gln Leu Ala Ala Leu Arg Pro Glu Val
370 375 380
Ala Glu Gln Thr Glu Thr Met Glu Gln Ala Met Arg Leu Leu Ala Pro
385 390 395 400
Arg Val Gln Pro Gly Asp Met Val Leu Leu Ser Pro Ala Cys Ala Ser
405 410 415
Leu Asp Gln Phe Lys Asn Phe Glu Gln Arg Gly Asn Glu Phe Ala Arg
420 425 430
Leu Ala Lys Glu Leu Gly
435

<210> 453
<211> 414
<212> PRT

<213> Escherichia coli

<400> 453
Met Arg Leu Ser Leu Pro Arg Leu Lys Met Pro Arg Leu Pro Gly Phe
1   5   10   15
Ser Ile Leu Val Trp Ile Ser Thr Ala Leu Lys Gly Trp Val Met Gly
            20   25   30
Ser Arg Glu Lys Asp Thr Asp Ser Leu Ile Met Tyr Asp Arg Thr Leu
        35   40   45
Leu Trp Leu Thr Phe Gly Leu Ala Ala Ile Gly Phe Ile Met Val Thr
    50   55   60
Ser Ala Ser Met Pro Ile Gly Gln Arg Leu Thr Asn Asp Pro Phe Phe
65   70   75   80
Phe Ala Lys Arg Asp Gly Val Tyr Leu Ile Leu Ala Phe Ile Leu Ala
                85   90   95
Ile Ile Thr Leu Arg Leu Pro Met Glu Phe Trp Gln Arg Tyr Ser Ala
            100   105   110
Thr Met Leu Leu Gly Ser Ile Ile Leu Leu Met Ile Val Leu Val Val
        115   120   125
Gly Ser Ser Val Lys Gly Ala Ser Arg Trp Ile Asp Leu Gly Leu Leu
    130   135   140
Arg Ile Gln Pro Ala Glu Leu Thr Lys Leu Ser Leu Phe Cys Tyr Ile
145   150   155   160
Ala Asn Tyr Leu Val Arg Lys Gly Asp Glu Val Arg Asn Asn Leu Arg
                165   170   175
Gly Phe Leu Lys Pro Met Gly Val Ile Leu Val Leu Ala Val Leu Leu
            180   185   190
Leu Ala Gln Pro Asp Leu Gly Thr Val Val Val Leu Phe Val Thr Thr
        195   200   205
Leu Ala Met Leu Phe Leu Ala Gly Ala Lys Leu Trp Gln Phe Ile Ala
    210   215   220
Ile Ile Gly Met Gly Ile Ser Ala Val Val Leu Leu Ile Leu Ala Glu
225   230   235   240
Pro Tyr Arg Ile Arg Arg Val Thr Ala Phe Trp Asn Pro Trp Glu Asp
                245   250   255
Pro Phe Gly Ser Gly Tyr Gln Leu Thr Gln Ser Leu Met Ala Phe Gly
            260   265   270
Arg Gly Glu Leu Trp Gly Gln Gly Leu Gly Asn Ser Val Gln Lys Leu
        275   280   285
Glu Tyr Leu Pro Glu Ala His Thr Asp Phe Ile Phe Ala Ile Ile Gly
    290   295   300
Glu Glu Leu Gly Tyr Val Gly Val Val Leu Ala Leu Leu Met Val Phe
305   310   315   320
Phe Val Ala Phe Arg Ala Met Ser Ile Gly Arg Lys Ala Leu Glu Ile
                325   330   335
Asp His Arg Phe Ser Gly Phe Leu Ala Cys Ser Ile Gly Ile Trp Phe
            340   345   350
Ser Phe Gln Ala Leu Val Asn Val Gly Ala Ala Ala Gly Met Leu Pro
        355   360   365
Thr Lys Gly Leu Thr Leu Pro Leu Ile Ser Tyr Gly Gly Ser Ser Leu
    370   375   380
Leu Ile Met Ser Thr Ala Ile Met Met Leu Leu Arg Ile Asp Tyr Glu
385   390   395   400
Thr Arg Leu Glu Lys Ala Gln Ala Phe Val Arg Gly Ser Arg
                405   410

<210> 454
<211> 355
<212> PRT
<213> Escherichia coli

<400> 454
Met Ser Gly Gln Gly Lys Arg Leu Met Val Met Ala Gly Gly Thr Gly
1 5 10 15
Gly His Val Phe Pro Gly Leu Ala Val Ala His His Leu Met Ala Gln
20 25 30
Gly Trp Gln Val Arg Trp Leu Gly Thr Ala Asp Arg Met Glu Ala Asp
35 40 45
Leu Val Pro Lys His Gly Ile Glu Ile Asp Phe Ile Arg Ile Ser Gly
50 55 60
Leu Arg Gly Lys Gly Ile Lys Ala Leu Ile Ala Ala Pro Leu Arg Ile
65 70 75 80
Phe Asn Ala Trp Arg Gln Ala Arg Ala Ile Met Lys Ala Tyr Lys Pro
85 90 95
Asp Val Val Leu Gly Met Gly Gly Tyr Val Ser Gly Pro Gly Gly Leu
100 105 110
Ala Ala Trp Ser Leu Gly Ile Pro Val Val Leu His Glu Gln Asn Gly
115 120 125
Ile Ala Gly Leu Thr Asn Lys Trp Leu Ala Lys Ile Ala Thr Lys Val
130 135 140
Met Gln Ala Phe Pro Gly Ala Phe Pro Asn Ala Glu Val Val Gly Asn
145 150 155 160
Pro Val Arg Thr Asp Val Leu Ala Leu Pro Leu Pro Gln Gln Arg Leu
165 170 175
Ala Gly Arg Glu Gly Pro Val Arg Val Leu Val Val Gly Gly Ser Gln
180 185 190
Gly Ala Arg Ile Leu Asn Gln Thr Met Pro Gln Val Ala Ala Lys Leu
195 200 205
Gly Asp Ser Val Thr Ile Trp His Gln Ser Gly Lys Gly Ser Gln Gln
210 215 220
Ser Val Glu Gln Ala Tyr Ala Glu Ala Gly Gln Pro Gln His Lys Val
225 230 235 240
Thr Glu Phe Ile Asp Asp Met Ala Ala Ala Tyr Ala Trp Ala Asp Val
245 250 255
Val Val Cys Arg Ser Gly Ala Leu Thr Val Ser Glu Ile Ala Ala Ala
260 265 270
Gly Leu Pro Ala Leu Phe Val Pro Phe Gln His Lys Asp Arg Gln Gln
275 280 285
Tyr Trp Asn Ala Leu Pro Leu Glu Lys Ala Gly Ala Ala Lys Ile Ile
290 295 300
Glu Gln Pro Gln Leu Ser Val Asp Ala Val Ala Asn Thr Leu Ala Gly
305 310 315 320
Trp Ser Arg Glu Thr Leu Leu Thr Met Ala Glu Arg Ala Arg Ala Ala
325 330 335
Ser Ile Pro Asp Ala Thr Glu Arg Val Ala Asn Glu Val Ser Arg Val
340 345 350
Ala Arg Ala
355

<210> 455
<211> 491
<212> PRT
<213> Escherichia coli

<400> 455
Met Asn Thr Gln Gln Leu Ala Lys Leu Arg Ser Ile Val Pro Glu Met
1 5 10 15
Arg Arg Val Arg His Ile His Phe Val Gly Ile Gly Gly Ala Gly Met
20 25 30
Gly Gly Ile Ala Glu Val Leu Ala Asn Glu Gly Tyr Gln Ile Ser Gly

```
        35                  40                  45
Ser Asp Leu Ala Pro Asn Pro Val Thr Gln Gln Leu Met Asn Leu Gly
    50                  55                  60
Ala Thr Ile Tyr Phe Asn His Arg Pro Glu Asn Val Arg Asp Ala Ser
65                  70                  75                  80
Val Val Val Val Ser Ser Ala Ile Ser Ala Asp Asn Pro Glu Ile Val
                85                  90                  95
Ala Ala His Glu Ala Arg Ile Pro Val Ile Arg Arg Ala Glu Met Leu
            100                 105                 110
Ala Glu Leu Met Arg Phe Arg His Gly Ile Ala Ile Ala Gly Thr His
        115                 120                 125
Gly Lys Thr Thr Thr Thr Ala Met Val Ser Ser Ile Tyr Ala Glu Ala
    130                 135                 140
Gly Leu Asp Pro Thr Phe Val Asn Gly Gly Leu Val Lys Ala Ala Gly
145                 150                 155                 160
Val His Ala Arg Leu Gly His Gly Arg Tyr Leu Ile Ala Glu Ala Asp
                165                 170                 175
Glu Ser Asp Ala Ser Phe Leu His Leu Gln Pro Met Val Ala Ile Val
            180                 185                 190
Thr Asn Ile Glu Ala Asp His Met Asp Thr Tyr Gln Gly Asp Phe Glu
        195                 200                 205
Asn Leu Lys Gln Thr Phe Ile Asn Phe Leu His Asn Leu Pro Phe Tyr
    210                 215                 220
Gly Arg Ala Val Met Cys Val Asp Asp Pro Val Ile Arg Glu Leu Leu
225                 230                 235                 240
Pro Arg Val Gly Arg Gln Thr Thr Thr Tyr Gly Phe Ser Glu Asp Ala
                245                 250                 255
Asp Val Arg Val Glu Asp Tyr Gln Gln Ile Gly Pro Gln Gly His Phe
            260                 265                 270
Thr Leu Leu Arg Gln Asp Lys Glu Pro Met Arg Val Thr Leu Asn Ala
        275                 280                 285
Pro Gly Arg His Asn Ala Leu Asn Ala Ala Ala Ala Val Ala Val Ala
    290                 295                 300
Thr Glu Glu Gly Ile Asp Asp Glu Ala Ile Leu Arg Ala Leu Glu Ser
305                 310                 315                 320
Phe Gln Gly Thr Gly Arg Arg Phe Asp Phe Leu Gly Glu Phe Pro Leu
                325                 330                 335
Glu Pro Val Asn Gly Lys Ser Gly Thr Ala Met Leu Val Asp Asp Tyr
            340                 345                 350
Gly His His Pro Thr Glu Val Asp Ala Thr Ile Lys Ala Ala Arg Ala
        355                 360                 365
Gly Trp Pro Asp Lys Asn Leu Val Met Leu Phe Gln Pro His Arg Phe
    370                 375                 380
Thr Arg Thr Arg Asp Leu Tyr Asp Asp Phe Ala Asn Val Leu Thr Gln
385                 390                 395                 400
Val Asp Thr Leu Leu Met Leu Glu Val Tyr Pro Ala Gly Glu Ala Pro
                405                 410                 415
Ile Pro Gly Ala Asp Ser Arg Ser Leu Cys Arg Thr Ile Arg Gly Arg
            420                 425                 430
Gly Lys Ile Asp Pro Ile Leu Val Pro Asp Pro Ala Arg Val Ala Glu
        435                 440                 445
Met Leu Ala Pro Val Leu Thr Gly Asn Asp Leu Ile Leu Val Gln Gly
    450                 455                 460
Ala Gly Asn Ile Gly Lys Ile Ala Arg Ser Leu Ala Glu Ile Lys Leu
465                 470                 475                 480
Lys Pro Gln Thr Pro Glu Glu Glu Gln His Asp
                485                 490
```

<210> 456
<211> 306

<212> PRT
<213> Escherichia coli

<400> 456
Met Thr Asp Lys Ile Ala Val Leu Leu Gly Gly Thr Ser Ala Glu Arg
1 5 10 15
Glu Val Ser Leu Asn Ser Gly Ala Ala Val Leu Ala Gly Leu Arg Glu
20 25 30
Gly Gly Ile Asp Ala Tyr Pro Val Asp Pro Lys Glu Val Asp Val Thr
35 40 45
Gln Leu Lys Ser Met Gly Phe Gln Lys Val Phe Ile Ala Leu His Gly
50 55 60
Arg Gly Gly Glu Asp Gly Thr Leu Gln Gly Met Leu Glu Leu Met Gly
65 70 75 80
Leu Pro Tyr Thr Gly Ser Gly Val Met Ala Ser Ala Leu Ser Met Asp
85 90 95
Lys Leu Arg Ser Lys Leu Leu Trp Gln Gly Ala Gly Leu Pro Val Ala
100 105 110
Pro Trp Val Ala Leu Thr Arg Ala Glu Phe Glu Lys Gly Leu Ser Asp
115 120 125
Lys Gln Leu Ala Glu Ile Ser Ala Leu Gly Leu Pro Val Ile Val Lys
130 135 140
Pro Ser Arg Glu Gly Ser Ser Val Gly Met Ser Lys Val Val Ala Glu
145 150 155 160
Asn Ala Leu Gln Asp Ala Leu Arg Leu Ala Phe Gln His Asp Glu Glu
165 170 175
Val Leu Ile Glu Lys Trp Leu Ser Gly Pro Glu Phe Thr Val Ala Ile
180 185 190
Leu Gly Glu Glu Ile Leu Pro Ser Ile Arg Ile Gln Pro Ser Gly Thr
195 200 205
Phe Tyr Asp Tyr Glu Ala Lys Tyr Leu Ser Asp Glu Thr Gln Tyr Phe
210 215 220
Cys Pro Ala Gly Leu Glu Ala Ser Gln Glu Ala Asn Leu Gln Ala Leu
225 230 235 240
Val Leu Lys Ala Trp Thr Thr Leu Gly Cys Lys Gly Trp Gly Arg Ile
245 250 255
Asp Val Met Leu Asp Ser Asp Gly Gln Phe Tyr Leu Leu Glu Ala Asn
260 265 270
Thr Ser Pro Gly Met Thr Ser His Ser Leu Val Pro Met Ala Ala Arg
275 280 285
Gln Ala Gly Met Ser Phe Ser Gln Leu Val Val Arg Ile Leu Glu Leu
290 295 300
Ala Asp
305

<210> 457
<211> 201
<212> PRT
<213> Escherichia coli

<400> 457
Met Ala Leu His Asp Glu Asn Val Val Trp His Ser His Pro Val Thr
1 5 10 15
Val Gln Gln Arg Glu Leu His His Gly His Arg Gly Val Val Leu Trp
20 25 30
Phe Thr Gly Leu Ser Gly Ser Gly Lys Ser Thr Val Ala Gly Ala Leu
35 40 45
Glu Glu Ala Leu His Lys Leu Gly Val Ser Thr Tyr Leu Leu Asp Gly
50 55 60
Asp Asn Val Arg His Gly Leu Cys Ser Asp Leu Gly Phe Ser Asp Ala

```
65                  70                  75                  80
Asp Arg Lys Glu Asn Ile Arg Arg Val Gly Glu Val Ala Asn Leu Met
                85                  90                  95
Val Glu Ala Gly Leu Val Val Leu Thr Ala Phe Ile Ser Pro His Arg
            100                 105                 110
Ala Glu Arg Gln Met Val Arg Glu Arg Val Gly Glu Gly Arg Phe Ile
        115                 120                 125
Glu Val Phe Val Asp Thr Pro Leu Ala Ile Cys Glu Ala Arg Asp Pro
    130                 135                 140
Lys Gly Leu Tyr Lys Lys Ala Arg Ala Gly Glu Leu Arg Asn Phe Thr
145                 150                 155                 160
Gly Ile Asp Ser Val Tyr Glu Ala Pro Glu Ser Ala Glu Ile His Leu
                165                 170                 175
Asn Gly Glu Gln Leu Val Thr Asn Leu Val Gln Gln Leu Leu Asp Leu
            180                 185                 190
Leu Arg Gln Asn Asp Ile Ile Arg Ser
        195                 200


<210> 458
<211> 475
<212> PRT
<213> Escherichia coli

<400> 458
Met Asn Thr Ala Leu Ala Gln Gln Ile Ala Asn Glu Gly Gly Val Glu
 1               5                  10                  15
Ala Trp Met Ile Ala Gln Gln His Lys Ser Leu Leu Arg Phe Leu Thr
            20                  25                  30
Cys Gly Ser Val Asp Asp Gly Lys Ser Thr Leu Ile Gly Arg Leu Leu
        35                  40                  45
His Asp Thr Arg Gln Ile Tyr Glu Asp Gln Leu Ser Ser Leu His Asn
    50                  55                  60
Asp Ser Lys Arg His Gly Thr Gln Gly Glu Lys Leu Asp Leu Ala Leu
65                  70                  75                  80
Leu Val Asp Gly Leu Gln Ala Glu Arg Glu Gln Gly Ile Thr Ile Asp
                85                  90                  95
Val Ala Tyr Arg Tyr Phe Ser Thr Glu Lys Arg Lys Phe Ile Ile Ala
            100                 105                 110
Asp Thr Pro Gly His Glu Gln Tyr Thr Arg Asn Met Ala Thr Gly Ala
        115                 120                 125
Ser Thr Cys Glu Leu Ala Ile Leu Leu Ile Asp Ala Arg Lys Gly Val
    130                 135                 140
Leu Asp Gln Thr Arg Arg His Ser Phe Ile Ser Thr Leu Leu Gly Ile
145                 150                 155                 160
Lys His Leu Val Val Ala Ile Asn Lys Met Asp Leu Val Asp Tyr Ser
                165                 170                 175
Glu Glu Thr Phe Thr Arg Ile Arg Glu Asp Tyr Leu Thr Phe Ala Gly
            180                 185                 190
Gln Leu Pro Gly Asn Leu Asp Ile Arg Phe Val Pro Leu Ser Ala Leu
        195                 200                 205
Glu Gly Asp Asn Val Ala Ser Gln Ser Glu Ser Met Pro Trp Tyr Ser
    210                 215                 220
Gly Pro Thr Leu Leu Glu Val Leu Glu Thr Val Glu Ile Gln Arg Val
225                 230                 235                 240
Val Asp Ala Gln Pro Met Arg Phe Pro Val Gln Tyr Val Asn Arg Pro
                245                 250                 255
Asn Leu Asp Phe Arg Gly Tyr Ala Gly Thr Leu Ala Ser Gly Arg Val
            260                 265                 270
Glu Val Gly Gln Arg Val Lys Val Leu Pro Ser Gly Val Glu Ser Asn
        275                 280                 285
```

Val Ala Arg Ile Val Thr Phe Asp Gly Asp Arg Glu Glu Ala Phe Ala
290 295 300
Gly Glu Ala Ile Thr Leu Val Leu Thr Asp Glu Ile Asp Ile Ser Arg
305 310 315 320
Gly Asp Leu Leu Leu Ala Ala Asp Glu Ala Leu Pro Ala Val Gln Ser
325 330 335
Ala Ser Val Asp Val Val Trp Met Ala Glu Gln Pro Leu Ser Pro Gly
340 345 350
Gln Ser Tyr Asp Ile Lys Ile Ala Gly Lys Lys Thr Arg Ala Arg Val
355 360 365
Asp Gly Ile Arg Tyr Gln Val Asp Ile Asn Asn Leu Thr Gln Arg Glu
370 375 380
Val Glu Asn Leu Pro Leu Asn Gly Ile Gly Leu Val Asp Leu Thr Phe
385 390 395 400
Asp Glu Pro Leu Val Leu Asp Arg Tyr Gln Gln Asn Pro Val Thr Gly
405 410 415
Gly Leu Ile Phe Ile Asp Arg Leu Ser Asn Val Thr Val Gly Ala Gly
420 425 430
Met Val His Glu Pro Val Ser Gln Ala Thr Ala Ala Pro Ser Glu Phe
435 440 445
Ser Ala Phe Glu Leu Glu Leu Asn Ala Leu Val Arg Arg His Phe Pro
450 455 460
His Trp Gly Ala Arg Asp Leu Leu Gly Asp Lys
465 470 475

<210> 459
<211> 127
<212> PRT
<213> Escherichia coli

<400> 459
Met Arg His Arg Lys Ser Gly Arg Gln Leu Asn Arg Asn Ser Ser His
1 5 10 15
Arg Gln Ala Met Phe Arg Asn Met Ala Gly Ser Leu Val Arg His Glu
20 25 30
Ile Ile Lys Thr Thr Leu Pro Lys Ala Lys Glu Leu Arg Arg Val Val
35 40 45
Glu Pro Leu Ile Thr Leu Ala Lys Thr Asp Ser Val Ala Asn Arg Arg
50 55 60
Leu Ala Phe Ala Arg Thr Arg Asp Asn Glu Ile Val Ala Lys Leu Phe
65 70 75 80
Asn Glu Leu Gly Pro Arg Phe Ala Ser Arg Ala Gly Gly Tyr Thr Arg
85 90 95
Ile Leu Lys Cys Gly Phe Arg Ala Gly Asp Asn Ala Pro Met Ala Tyr
100 105 110
Ile Glu Leu Val Asp Arg Ser Glu Lys Ala Glu Ala Ala Ala Glu
115 120 125

<210> 460
<211> 329
<212> PRT
<213> Escherichia coli

<400> 460
Met Gln Gly Ser Val Thr Glu Phe Leu Lys Pro Arg Leu Val Asp Ile
1 5 10 15
Glu Gln Val Ser Ser Thr His Ala Lys Val Thr Leu Glu Pro Leu Glu
20 25 30
Arg Gly Phe Gly His Thr Leu Gly Asn Ala Leu Arg Arg Ile Leu Leu

```
        35                  40                  45
Ser Ser Met Pro Gly Cys Ala Val Thr Glu Val Glu Ile Asp Gly Val
    50                  55                  60
Leu His Glu Tyr Ser Thr Lys Glu Gly Val Gln Glu Asp Ile Leu Glu
65                  70                  75                  80
Ile Leu Leu Asn Leu Lys Gly Leu Ala Val Arg Val Gln Gly Lys Asp
                85                  90                  95
Glu Val Ile Leu Thr Leu Asn Lys Ser Gly Ile Gly Pro Val Thr Ala
            100                 105                 110
Ala Asp Ile Thr His Asp Gly Asp Val Glu Ile Val Lys Pro Gln His
        115                 120                 125
Val Ile Cys His Leu Thr Asp Glu Asn Ala Ser Ile Ser Met Arg Ile
    130                 135                 140
Lys Val Gln Arg Gly Arg Gly Tyr Val Pro Ala Ser Thr Arg Ile His
145                 150                 155                 160
Ser Glu Glu Asp Glu Arg Pro Ile Gly Arg Leu Leu Val Asp Ala Cys
                165                 170                 175
Tyr Ser Pro Val Glu Arg Ile Ala Tyr Asn Val Glu Ala Ala Arg Val
            180                 185                 190
Glu Gln Arg Thr Asp Leu Asp Lys Leu Val Ile Glu Met Glu Thr Asn
        195                 200                 205
Gly Thr Ile Asp Pro Glu Glu Ala Ile Arg Arg Ala Ala Thr Ile Leu
    210                 215                 220
Ala Glu Gln Leu Glu Ala Phe Val Asp Leu Arg Asp Val Arg Gln Pro
225                 230                 235                 240
Glu Val Lys Glu Glu Lys Pro Glu Phe Asp Pro Ile Leu Leu Arg Pro
                245                 250                 255
Val Asp Asp Leu Glu Leu Thr Val Arg Ser Ala Asn Cys Leu Lys Ala
            260                 265                 270
Glu Ala Ile His Tyr Ile Gly Asp Leu Val Gln Arg Thr Glu Val Glu
        275                 280                 285
Leu Leu Lys Thr Pro Asn Leu Gly Lys Lys Ser Leu Thr Glu Ile Lys
    290                 295                 300
Asp Val Leu Ala Ser Arg Gly Leu Ser Leu Gly Met Arg Leu Glu Asn
305                 310                 315                 320
Trp Pro Pro Ala Ser Ile Ala Asp Glu
                325
```

```
<210> 461
<211> 206
<212> PRT
<213> Escherichia coli

<400> 461
Met Ala Arg Tyr Leu Gly Pro Lys Leu Lys Leu Ser Arg Arg Glu Gly
 1               5                  10                  15
Thr Asp Leu Phe Leu Lys Ser Gly Val Arg Ala Ile Asp Thr Lys Cys
            20                  25                  30
Lys Ile Glu Gln Ala Pro Gly Gln His Gly Ala Arg Lys Pro Arg Leu
        35                  40                  45
Ser Asp Tyr Gly Val Gln Leu Arg Glu Lys Gln Lys Val Arg Arg Ile
    50                  55                  60
Tyr Gly Val Leu Glu Arg Gln Phe Arg Asn Tyr Tyr Lys Glu Ala Ala
65                  70                  75                  80
Arg Leu Lys Gly Asn Thr Gly Glu Asn Leu Leu Ala Leu Leu Glu Gly
                85                  90                  95
Arg Leu Asp Asn Val Val Tyr Arg Met Gly Phe Gly Ala Thr Arg Ala
            100                 105                 110
Glu Ala Arg Gln Leu Val Ser His Lys Ala Ile Met Val Asn Gly Arg
        115                 120                 125
```

Val Val Asn Ile Ala Ser Tyr Gln Val Ser Pro Asn Asp Val Val Ser
130 135 140
Ile Arg Glu Lys Ala Lys Lys Gln Ser Arg Val Lys Ala Ala Leu Glu
145 150 155 160
Leu Ala Glu Gln Arg Glu Lys Pro Thr Trp Leu Glu Val Asp Ala Gly
165 170 175
Lys Met Glu Gly Thr Phe Lys Arg Lys Pro Glu Arg Ser Asp Leu Ser
180 185 190
Ala Asp Ile Asn Glu His Leu Ile Val Glu Leu Tyr Ser Lys
195 200 205

<210> 462
<211> 129
<212> PRT
<213> Escherichia coli

<400> 462
Met Ala Lys Ala Pro Ile Arg Ala Arg Lys Arg Val Arg Lys Gln Val
1 5 10 15
Ser Asp Gly Val Ala His Ile His Ala Ser Phe Asn Asn Thr Ile Val
20 25 30
Thr Ile Thr Asp Arg Gln Gly Asn Ala Leu Gly Trp Ala Thr Ala Gly
35 40 45
Gly Ser Gly Phe Arg Gly Ser Arg Lys Ser Thr Pro Phe Ala Ala Gln
50 55 60
Val Ala Ala Glu Arg Cys Ala Asp Ala Val Lys Glu Tyr Gly Ile Lys
65 70 75 80
Asn Leu Glu Val Met Val Lys Gly Pro Gly Pro Gly Arg Glu Ser Thr
85 90 95
Ile Arg Ala Leu Asn Ala Ala Gly Phe Arg Ile Thr Asn Ile Thr Asp
100 105 110
Val Thr Pro Ile Pro His Asn Gly Cys Arg Pro Pro Lys Lys Arg Arg
115 120 125
Val

<210> 463
<211> 118
<212> PRT
<213> Escherichia coli

<400> 463
Met Ala Arg Ile Ala Gly Ile Asn Ile Pro Asp His Lys His Ala Val
1 5 10 15
Ile Ala Leu Thr Ser Ile Tyr Gly Val Gly Lys Thr Arg Ser Lys Ala
20 25 30
Ile Leu Ala Ala Ala Gly Ile Ala Glu Asp Val Lys Ile Ser Glu Leu
35 40 45
Ser Glu Gly Gln Ile Asp Thr Leu Arg Asp Glu Val Ala Lys Phe Val
50 55 60
Val Glu Gly Asp Leu Arg Arg Glu Ile Ser Met Ser Ile Lys Arg Leu
65 70 75 80
Met Asp Leu Gly Cys Tyr Arg Gly Leu Arg His Arg Arg Gly Leu Pro
85 90 95
Val Arg Gly Gln Arg Thr Lys Thr Asn Ala Arg Thr Arg Lys Gly Pro
100 105 110
Arg Lys Pro Ile Lys Lys
115

<210> 464
<211> 460
<212> PRT
<213> Escherichia coli

<400> 464
Met Thr Gln Leu Thr Met Lys Asp Lys Ile Gly Tyr Gly Leu Gly Asp
1 5 10 15
Thr Ala Cys Gly Phe Val Trp Gln Ala Thr Met Phe Leu Leu Ala Tyr
20 25 30
Phe Tyr Thr Asp Val Phe Gly Leu Ser Ala Gly Ile Met Gly Thr Leu
35 40 45
Phe Leu Val Ser Arg Val Leu Asp Ala Val Thr Asp Pro Leu Met Gly
50 55 60
Leu Leu Val Asp Arg Thr Arg Thr Arg His Gly Gln Phe Arg Pro Phe
65 70 75 80
Leu Leu Trp Gly Ala Ile Pro Phe Gly Ile Val Cys Val Leu Thr Phe
85 90 95
Tyr Thr Pro Asp Phe Ser Ala Gln Gly Lys Ile Ile Tyr Ala Cys Val
100 105 110
Thr Tyr Ile Leu Leu Thr Leu Val Tyr Thr Phe Val Asn Val Pro Tyr
115 120 125
Cys Ala Met Pro Gly Val Ile Thr Ala Asp Pro Lys Glu Arg His Ala
130 135 140
Leu Gln Ser Trp Arg Phe Phe Leu Ala Ala Ala Gly Ser Leu Ala Ile
145 150 155 160
Ser Gly Ile Ala Leu Pro Leu Val Ser Ile Ile Gly Lys Gly Asp Glu
165 170 175
Gln Val Gly Tyr Phe Gly Ala Met Cys Val Leu Gly Leu Ser Gly Val
180 185 190
Val Leu Leu Tyr Val Cys Phe Phe Thr Thr Lys Glu Arg Tyr Thr Phe
195 200 205
Glu Val Gln Pro Gly Ser Ser Val Ala Lys Asp Leu Lys Leu Leu Leu
210 215 220
Gly Asn Ser Gln Trp Arg Ile Met Cys Ala Phe Lys Met Met Ala Thr
225 230 235 240
Cys Ser Asn Val Val Arg Gly Gly Ala Thr Leu Tyr Phe Val Lys Tyr
245 250 255
Val Met Asp His Pro Glu Leu Ala Thr Gln Phe Leu Leu Tyr Gly Ser
260 265 270
Leu Ala Thr Met Phe Gly Ser Leu Cys Ser Ser Arg Leu Leu Gly Arg
275 280 285
Phe Asp Arg Val Thr Ala Phe Lys Trp Ile Ile Val Ala Tyr Ser Leu
290 295 300
Ile Ser Leu Leu Ile Phe Val Thr Pro Ala Glu His Ile Ala Leu Ile
305 310 315 320
Phe Ala Leu Asn Ile Leu Phe Leu Phe Val Phe Asn Thr Thr Thr Pro
325 330 335
Leu Gln Trp Leu Met Ala Ser Asp Val Val Asp Tyr Glu Glu Ser Arg
340 345 350
Ser Gly Arg Arg Leu Asp Gly Leu Val Phe Ser Thr Tyr Leu Phe Ser
355 360 365
Leu Lys Ile Gly Leu Ala Ile Gly Gly Ala Val Val Gly Trp Ile Leu
370 375 380
Ala Tyr Val Asn Tyr Ser Ala Ser Ser Ser Val Gln Pro Val Glu Val
385 390 395 400
Leu Thr Thr Ile Lys Ile Leu Phe Cys Val Val Pro Val Val Leu Tyr
405 410 415
Ala Gly Met Phe Ile Met Leu Ser Leu Tyr Lys Leu Thr Asp Ala Arg
420 425 430

Val Glu Ala Ile Ser Arg Gln Leu Ile Lys His Arg Ala Ala Gln Gly
435 440 445
Glu Ala Val Pro Asp Ala Ala Thr Ala Ala Ser His
450 455 460

<210> 465
<211> 536
<212> PRT
<213> Escherichia coli

<400> 465
Met Glu Ile Thr Asn Pro Ile Leu Thr Gly Phe Asn Pro Asp Pro Ser
1 5 10 15
Leu Cys Arg Gln Gly Glu Asp Tyr Tyr Ile Ala Thr Ser Thr Phe Glu
20 25 30
Trp Phe Pro Gly Val Arg Ile Tyr His Ser Arg Asp Leu Lys Asn Trp
35 40 45
Ser Leu Val Ser Thr Pro Leu Asp Arg Val Ser Met Leu Asp Met Lys
50 55 60
Gly Asn Pro Asp Ser Gly Gly Ile Trp Ala Pro Cys Leu Ser Tyr Ala
65 70 75 80
Asp Gly Lys Phe Trp Leu Leu Tyr Thr Asp Val Lys Ile Val Asp Ser
85 90 95
Pro Trp Lys Asn Gly Arg Asn Phe Leu Val Thr Ala Pro Ser Ile Glu
100 105 110
Gly Pro Trp Ser Glu Pro Ile Pro Met Gly Asn Gly Gly Phe Asp Pro
115 120 125
Ser Leu Phe His Asp Asp Asp Gly Arg Lys Tyr Tyr Ile Tyr Arg Pro
130 135 140
Trp Gly Pro Arg His His Ser Asn Pro His Asn Thr Ile Val Leu Gln
145 150 155 160
Ala Phe Asp Pro Gln Thr Gly Thr Leu Ser Pro Glu Arg Lys Thr Leu
165 170 175
Phe Thr Gly Thr Pro Leu Cys Tyr Thr Glu Gly Ala His Leu Tyr Arg
180 185 190
His Ala Gly Trp Tyr Tyr Leu Met Ala Ala Glu Gly Gly Thr Ser Tyr
195 200 205
Glu His Ala Val Val Val Leu Arg Ser Lys Asn Ile Asp Gly Pro Tyr
210 215 220
Glu Leu His Pro Asp Val Thr Met Met Thr Ser Trp His Leu Pro Glu
225 230 235 240
Asn Pro Leu Gln Lys Ser Gly His Gly Ser Leu Leu Gln Thr His Thr
245 250 255
Gly Glu Trp Tyr Met Ala Tyr Leu Thr Ser Arg Pro Leu Arg Leu Pro
260 265 270
Gly Val Pro Leu Leu Ala Ser Gly Gly Arg Gly Tyr Cys Pro Leu Gly
275 280 285
Arg Glu Thr Gly Ile Ala Arg Ile Glu Trp Arg Asp Gly Trp Pro Tyr
290 295 300
Val Glu Gly Gly Lys His Ala Gln Leu Thr Val Lys Gly Pro Gln Val
305 310 315 320
Ala Glu Gln Pro Ala Ala Val Pro Gly Asn Trp Arg Asp Asp Phe Asp
325 330 335
Ala Ser Ser Leu Asp Pro Glu Leu Gln Thr Leu Arg Ile Pro Phe Asp
340 345 350
Asp Thr Leu Gly Ser Leu Thr Ala Arg Pro Gly Phe Leu Arg Leu Tyr
355 360 365
Gly Asn Asp Ser Leu Asn Ser Thr Phe Thr Gln Ser Thr Val Ala Arg
370 375 380
Arg Trp Gln His Phe Ala Phe Arg Ala Glu Thr Arg Met Glu Phe Ser

385 390 395 400
Pro Val His Phe Gln Gln Ser Ala Gly Leu Thr Cys Tyr Tyr Asn Ser
405 410 415
Lys Asn Trp Ser Tyr Cys Phe Val Asp Tyr Glu Glu Gly Gln Gly Arg
420 425 430
Thr Ile Lys Val Ile Gln Leu Asp His Asn Val Pro Ser Trp Pro Leu
435 440 445
His Glu Gln Pro Ile Pro Val Pro Glu His Ala Glu Ser Val Trp Leu
450 455 460
Arg Val Asp Val Asp Thr Leu Val Tyr Arg Tyr Ser Tyr Ser Phe Asp
465 470 475 480
Gly Glu Thr Trp His Thr Val Pro Val Thr Tyr Glu Ala Trp Lys Leu
485 490 495
Ser Asp Asp Tyr Ile Gly Gly Arg Gly Phe Phe Thr Gly Ala Phe Val
500 505 510
Gly Leu His Cys Glu Asp Ile Ser Gly Asp Gly Cys Tyr Ala Asp Phe
515 520 525
Asp Tyr Phe Thr Tyr Glu Pro Val
530 535

<210> 466
<211> 325
<212> PRT
<213> Escherichia coli

<400> 466
Met His Met Lys Lys Ile Ile Phe Ala Phe Ile Ile Leu Phe Val Phe
1 5 10 15
Leu Leu Pro Met Ile Ile Phe Tyr Gln Pro Trp Val Asn Ala Leu Pro
20 25 30
Ser Thr Pro Arg His Ala Ser Pro Glu Gln Leu Glu Lys Thr Val Arg
35 40 45
Tyr Leu Thr Gln Thr Val His Pro Arg Ser Ala Asp Asn Ile Asp Asn
50 55 60
Leu Asn Arg Ser Ala Glu Tyr Ile Lys Glu Val Phe Val Ser Ser Gly
65 70 75 80
Ala Arg Val Thr Ser Gln Asp Val Pro Ile Thr Gly Gly Pro Tyr Lys
85 90 95
Asn Ile Val Ala Asp Tyr Gly Pro Ala Asp Gly Pro Leu Ile Ile Ile
100 105 110
Gly Ala His Tyr Asp Ser Ala Ser Ser Tyr Glu Asn Asp Gln Leu Thr
115 120 125
Tyr Thr Pro Gly Ala Asp Asp Asn Ala Ser Gly Val Ala Gly Leu Leu
130 135 140
Glu Leu Ala Arg Leu Leu His Gln Gln Val Pro Lys Thr Gly Val Gln
145 150 155 160
Leu Val Ala Tyr Ala Ser Glu Glu Pro Pro Phe Phe Arg Ser Asp Glu
165 170 175
Met Gly Ser Ala Val His Ala Ala Ser Leu Glu Arg Pro Val Lys Leu
180 185 190
Met Ile Ala Leu Glu Met Ile Gly Tyr Tyr Asp Ser Ala Pro Gly Ser
195 200 205
Gln Asn Tyr Pro Tyr Pro Ala Met Ser Trp Leu Tyr Pro Asp Arg Gly
210 215 220
Asp Phe Ile Ala Val Val Gly Arg Ile Gln Asp Ile Asn Ala Val Arg
225 230 235 240
Gln Val Lys Ala Ala Leu Leu Ser Ser Gln Asp Leu Ser Val Tyr Ser
245 250 255
Met Asn Thr Pro Gly Phe Ile Pro Gly Ile Asp Phe Ser Asp His Leu
260 265 270

Asn Tyr Trp Gln His Asp Ile Pro Ala Ile Met Ile Thr Asp Thr Ala
        275   280   285
Phe Tyr Arg Asn Lys Gln Tyr His Leu Pro Gly Asp Thr Ala Asp Arg
    290   295   300
Leu Asn Tyr Gln Lys Met Ala Gln Val Val Asp Gly Val Ile Thr Leu
305   310   315   320
Leu Tyr Asn Ser Lys
                325

<210> 467
<211> 320
<212> PRT
<213> Escherichia coli

<400> 467
Met Met Ile Lys Thr Arg Phe Ser Arg Trp Leu Thr Phe Phe Thr Phe
 1   5   10   15
Ala Ala Ala Val Ala Leu Ala Leu Pro Ala Lys Ala Asn Thr Trp Pro
            20   25   30
Leu Pro Pro Ala Gly Ser Arg Leu Val Gly Glu Asn Lys Phe His Val
        35   40   45
Val Glu Asn Asp Gly Gly Ser Leu Glu Ala Ile Ala Lys Lys Tyr Asn
    50   55   60
Val Gly Phe Leu Ala Leu Leu Gln Ala Asn Pro Gly Val Asp Pro Tyr
65   70   75   80
Val Pro Arg Ala Gly Ser Val Leu Thr Ile Pro Leu Gln Thr Leu Leu
                85   90   95
Pro Asp Ala Pro Arg Glu Gly Ile Val Ile Asn Ile Ala Glu Leu Arg
            100   105   110
Leu Tyr Tyr Tyr Pro Pro Gly Lys Asn Ser Val Thr Val Tyr Pro Ile
        115   120   125
Gly Ile Gly Gln Leu Gly Gly Asp Thr Leu Thr Pro Thr Met Val Thr
    130   135   140
Thr Val Ser Asp Lys Arg Ala Asn Pro Thr Trp Thr Pro Thr Ala Asn
145   150   155   160
Ile Arg Ala Arg Tyr Lys Ala Gln Gly Ile Glu Leu Pro Ala Val Val
                165   170   175
Pro Ala Gly Leu Asp Asn Pro Met Gly His His Ala Ile Arg Leu Ala
            180   185   190
Ala Tyr Gly Gly Val Tyr Leu Leu His Gly Thr Asn Ala Asp Phe Gly
        195   200   205
Ile Gly Met Arg Val Ser Ser Gly Cys Ile Arg Leu Arg Asp Asp Asp
    210   215   220
Ile Lys Thr Leu Phe Ser Gln Val Thr Pro Gly Thr Lys Val Asn Ile
225   230   235   240
Ile Asn Thr Pro Ile Lys Val Ser Ala Glu Pro Asn Gly Ala Arg Leu
                245   250   255
Val Glu Val His Gln Pro Leu Ser Glu Lys Ile Asp Asp Asp Pro Gln
            260   265   270
Leu Leu Pro Ile Thr Leu Asn Ser Ala Met Gln Ser Phe Lys Asp Ala
        275   280   285
Ala Gln Thr Asp Ala Glu Val Met Gln His Val Met Asp Val Arg Ser
    290   295   300
Gly Met Pro Val Asp Val Arg Arg His Gln Val Ser Pro Gln Thr Leu
305   310   315   320

<210> 468
<211> 494
<212> PRT

<213> Escherichia coli

<400> 468

```
Met Val Ala Ile His Leu Leu Pro Val Ser Tyr Asn Ser Ala Thr Ser
 1               5                  10                  15
Thr Val Asn Ile Ser Ala Arg Ile Ile Pro Leu Leu Ile Ile His Gln
            20                  25                  30
Arg Tyr Lys Ile Pro Met Pro Lys Val Gln Ala Asp Gly Leu Pro Leu
        35                  40                  45
Pro Gln Arg Tyr Gly Ala Ile Leu Thr Ile Val Ile Gly Ile Ser Met
    50                  55                  60
Ala Val Leu Asp Gly Ala Ile Ala Asn Val Ala Leu Pro Thr Ile Ala
65                  70                  75                  80
Thr Asp Leu His Ala Thr Pro Ala Ser Ser Ile Trp Val Val Asn Ala
                85                  90                  95
Tyr Gln Ile Ala Ile Val Ile Ser Leu Leu Ser Phe Ser Phe Leu Gly
            100                 105                 110
Asp Met Phe Gly Tyr Arg Arg Ile Tyr Lys Cys Gly Leu Val Val Phe
        115                 120                 125
Leu Leu Ser Ser Leu Phe Cys Ala Leu Ser Asp Ser Leu Gln Met Leu
    130                 135                 140
Thr Leu Ala Arg Val Ile Gln Gly Phe Gly Gly Ala Ala Leu Met Ser
145                 150                 155                 160
Val Asn Thr Ala Leu Ile Arg Leu Ile Tyr Pro Gln Arg Phe Leu Gly
                165                 170                 175
Arg Gly Met Gly Ile Asn Ser Phe Ile Val Ala Val Ser Ser Ala Ala
            180                 185                 190
Gly Pro Thr Ile Ala Ala Ala Ile Leu Ser Ile Ala Ser Trp Lys Trp
        195                 200                 205
Leu Phe Leu Ile Asn Val Pro Leu Gly Ile Ile Ala Leu Leu Leu Ala
    210                 215                 220
Met Arg Phe Leu Pro Pro Asn Gly Ser Arg Ala Ser Lys Pro Arg Phe
225                 230                 235                 240
Asp Leu Pro Ser Ala Val Met Asn Ala Leu Thr Phe Gly Leu Leu Ile
                245                 250                 255
Thr Ala Leu Ser Gly Phe Ala Gln Gly Gln Ser Leu Thr Leu Ile Ala
            260                 265                 270
Ala Glu Leu Val Val Met Val Val Val Gly Ile Phe Phe Ile Arg Arg
        275                 280                 285
Gln Leu Ser Leu Pro Val Pro Leu Leu Pro Val Asp Leu Leu Arg Ile
    290                 295                 300
Pro Leu Phe Ser Leu Ser Ile Cys Thr Ser Val Cys Ser Phe Cys Ala
305                 310                 315                 320
Gln Met Leu Ala Met Val Ser Leu Pro Phe Tyr Leu Gln Thr Val Leu
                325                 330                 335
Gly Arg Ser Glu Val Glu Thr Gly Leu Leu Leu Thr Pro Trp Pro Leu
            340                 345                 350
Ala Thr Met Val Met Ala Pro Leu Ala Gly Tyr Leu Ile Glu Arg Val
        355                 360                 365
His Ala Gly Leu Leu Gly Ala Leu Gly Leu Phe Ile Met Ala Ala Gly
    370                 375                 380
Leu Phe Ser Leu Val Leu Leu Pro Ala Ser Pro Ala Asp Ile Asn Ile
385                 390                 395                 400
Ile Trp Pro Met Ile Leu Cys Gly Ala Gly Phe Gly Leu Phe Gln Ser
                405                 410                 415
Pro Asn Asn His Thr Ile Ile Thr Ser Ala Pro Arg Glu Arg Ser Gly
            420                 425                 430
Gly Ala Ser Gly Met Leu Gly Thr Ala Arg Leu Leu Gly Gln Ser Ser
        435                 440                 445
Gly Ala Ala Leu Val Ala Leu Met Leu Asn Gln Phe Gly Asp Asn Gly
    450                 455                 460
```

Thr His Val Ser Leu Met Ala Ala Ala Ile Leu Ala Val Ile Ala Ala
465 470 475 480
Cys Val Ser Gly Leu Arg Ile Thr Gln Pro Arg Ser Arg Ala
485 490

<210> 469
<211> 477
<212> PRT
<213> Escherichia coli

<400> 469
Met Lys Val Thr Leu Pro Glu Phe Glu Arg Ala Gly Val Met Val Val
1 5 10 15
Gly Asp Val Met Leu Asp Arg Tyr Trp Tyr Gly Pro Thr Ser Arg Ile
20 25 30
Ser Pro Glu Ala Pro Val Pro Val Val Lys Val Asn Thr Ile Glu Glu
35 40 45
Arg Pro Gly Gly Ala Ala Asn Val Ala Met Asn Ile Ala Ser Leu Gly
50 55 60
Ala Asn Ala Arg Leu Val Gly Leu Thr Gly Ile Asp Asp Ala Ala Arg
65 70 75 80
Ala Leu Ser Lys Ser Leu Ala Asp Val Asn Val Lys Cys Asp Phe Val
85 90 95
Ser Val Pro Thr His Pro Thr Ile Thr Lys Leu Arg Val Leu Ser Arg
100 105 110
Asn Gln Gln Leu Ile Arg Leu Asp Phe Glu Glu Gly Phe Glu Gly Val
115 120 125
Asp Pro Gln Pro Leu His Glu Arg Ile Asn Gln Ala Leu Ser Ser Ile
130 135 140
Gly Ala Leu Val Leu Ser Asp Tyr Ala Lys Gly Ala Leu Ala Ser Val
145 150 155 160
Gln Gln Met Ile Gln Leu Ala Arg Lys Ala Gly Val Pro Val Leu Ile
165 170 175
Asp Pro Lys Gly Thr Asp Phe Glu Arg Tyr Arg Gly Ala Thr Leu Leu
180 185 190
Thr Pro Asn Leu Ser Glu Phe Glu Ala Val Val Gly Lys Cys Lys Thr
195 200 205
Glu Glu Glu Ile Val Glu Arg Gly Met Lys Leu Ile Ala Asp Tyr Glu
210 215 220
Leu Ser Ala Leu Leu Val Thr Arg Ser Glu Gln Gly Met Ser Leu Leu
225 230 235 240
Gln Pro Gly Lys Ala Pro Leu His Met Pro Thr Gln Ala Gln Glu Val
245 250 255
Tyr Asp Val Thr Gly Ala Gly Asp Thr Val Ile Gly Val Leu Ala Ala
260 265 270
Thr Leu Ala Ala Gly Asn Ser Leu Glu Glu Ala Cys Phe Phe Ala Asn
275 280 285
Ala Ala Ala Gly Val Val Val Gly Lys Leu Gly Thr Ser Thr Val Ser
290 295 300
Pro Ile Glu Leu Glu Asn Ala Val Arg Gly Arg Ala Asp Thr Gly Phe
305 310 315 320
Gly Val Met Thr Glu Glu Glu Leu Lys Leu Ala Val Ala Ala Ala Arg
325 330 335
Lys Arg Gly Glu Lys Val Val Met Thr Asn Gly Val Phe Asp Ile Leu
340 345 350
His Ala Gly His Val Ser Tyr Leu Ala Asn Ala Arg Lys Leu Gly Asp
355 360 365
Arg Leu Ile Val Ala Val Asn Ser Asp Ala Ser Thr Lys Arg Leu Lys
370 375 380
Gly Asp Ser Arg Pro Val Asn Pro Leu Glu Gln Arg Met Ile Val Leu

385 390 395 400
Gly Ala Leu Glu Ala Val Asp Trp Val Val Ser Phe Glu Glu Asp Thr
405 410 415
Pro Gln Arg Leu Ile Ala Gly Ile Leu Pro Asp Leu Leu Val Lys Gly
420 425 430
Gly Asp Tyr Lys Pro Glu Glu Ile Ala Gly Ser Lys Glu Val Trp Ala
435 440 445
Asn Gly Gly Glu Val Leu Val Leu Asn Phe Glu Asp Gly Cys Ser Thr
450 455 460
Thr Asn Ile Ile Lys Lys Ile Gln Gln Asp Lys Lys Gly
465 470 475

<210> 470
<211> 946
<212> PRT
<213> Escherichia coli

<400> 470
Met Lys Pro Leu Ser Ser Pro Leu Gln Gln Tyr Trp Gln Thr Val Val
1 5 10 15
Glu Arg Leu Pro Glu Pro Leu Ala Glu Glu Ser Leu Ser Ala Gln Ala
20 25 30
Lys Ser Val Leu Thr Phe Ser Asp Phe Val Gln Asp Ser Val Ile Ala
35 40 45
His Pro Glu Trp Leu Thr Glu Leu Glu Ser Gln Pro Pro Gln Ala Asp
50 55 60
Glu Trp Gln His Tyr Ala Ala Trp Leu Gln Glu Ala Leu Cys Asn Val
65 70 75 80
Ser Asp Glu Ala Gly Leu Met Arg Glu Leu Arg Leu Phe Arg Arg Arg
85 90 95
Ile Met Val Arg Ile Ala Trp Ala Gln Thr Leu Ala Leu Val Thr Glu
100 105 110
Glu Ser Ile Leu Gln Gln Leu Ser Tyr Leu Ala Glu Thr Leu Ile Val
115 120 125
Ala Ala Arg Asp Trp Leu Tyr Asp Ala Cys Cys Arg Glu Trp Gly Thr
130 135 140
Pro Cys Asn Ala Gln Gly Glu Ala Gln Pro Leu Leu Ile Leu Gly Met
145 150 155 160
Gly Lys Leu Gly Gly Gly Glu Leu Asn Phe Ser Ser Asp Ile Asp Leu
165 170 175
Ile Phe Ala Trp Pro Glu His Gly Cys Thr Gln Gly Gly Arg Arg Glu
180 185 190
Leu Asp Asn Ala Gln Phe Phe Thr Arg Met Gly Gln Arg Leu Ile Lys
195 200 205
Val Leu Asp Gln Pro Thr Gln Asp Gly Phe Val Tyr Arg Val Asp Met
210 215 220
Arg Leu Arg Pro Phe Gly Glu Ser Gly Pro Leu Val Leu Ser Phe Ala
225 230 235 240
Ala Leu Glu Asp Tyr Tyr Gln Glu Gln Gly Arg Asp Trp Glu Arg Tyr
245 250 255
Ala Met Val Lys Ala Arg Ile Met Gly Asp Ser Glu Gly Val Tyr Ala
260 265 270
Asn Glu Leu Arg Ala Met Leu Arg Pro Phe Val Phe Arg Arg Tyr Ile
275 280 285
Asp Phe Ser Val Ile Gln Ser Leu Arg Asn Met Lys Gly Met Ile Ala
290 295 300
Arg Glu Val Arg Arg Arg Gly Leu Thr Asp Asn Ile Lys Leu Gly Ala
305 310 315 320
Gly Gly Ile Arg Glu Ile Glu Phe Ile Val Gln Val Phe Gln Leu Ile
325 330 335

Arg Gly Gly Arg Glu Pro Ser Leu Gln Ser Arg Ser Leu Leu Pro Thr
340 345 350
Leu Ser Ala Ile Ala Glu Leu His Leu Leu Ser Glu Asn Asp Ala Glu
355 360 365
Gln Leu Arg Val Ala Tyr Leu Phe Leu Arg Arg Leu Glu Asn Leu Leu
370 375 380
Gln Ser Ile Asn Asp Glu Gln Thr Gln Thr Leu Pro Ser Asp Glu Leu
385 390 395 400
Asn Arg Ala Arg Leu Ala Trp Ala Met Asp Phe Ala Asp Trp Pro Gln
405 410 415
Leu Thr Gly Ala Leu Thr Ala His Met Thr Asn Val Arg Arg Val Phe
420 425 430
Asn Glu Leu Ile Gly Asp Asp Glu Ser Glu Thr Gln Glu Glu Ser Leu
435 440 445
Ser Glu Gln Trp Arg Glu Leu Trp Gln Asp Ala Leu Gln Glu Asp Asp
450 455 460
Thr Thr Pro Val Leu Ala His Leu Ser Glu Asp Asp Arg Lys Gln Val
465 470 475 480
Leu Thr Leu Ile Ala Asp Phe Arg Lys Glu Leu Asp Lys Arg Thr Ile
485 490 495
Gly Pro Arg Gly Arg Gln Val Leu Asp His Leu Met Pro His Leu Leu
500 505 510
Ser Asp Val Cys Ala Arg Glu Asp Ala Ala Val Thr Leu Ser Arg Ile
515 520 525
Thr Ala Leu Leu Val Gly Ile Val Thr Arg Thr Thr Tyr Leu Glu Leu
530 535 540
Leu Ser Glu Phe Pro Ala Ala Leu Lys His Leu Ile Ser Leu Cys Ala
545 550 555 560
Ala Ser Pro Met Ile Ala Ser Gln Leu Ala Arg Tyr Pro Leu Leu Leu
565 570 575
Asp Glu Leu Leu Asp Pro Asn Thr Leu Tyr Gln Pro Thr Ala Thr Asp
580 585 590
Ala Tyr Arg Asp Glu Leu Arg Gln Tyr Leu Leu Arg Val Pro Glu Asp
595 600 605
Asp Glu Glu Gln Gln Leu Glu Ala Leu Arg Gln Phe Lys Gln Ala Gln
610 615 620
Leu Leu Arg Ile Ala Ala Ala Asp Ile Ala Gly Thr Leu Pro Val Met
625 630 635 640
Lys Val Ser Asp His Leu Thr Trp Leu Ala Glu Ala Met Ile Asp Ala
645 650 655
Val Val Gln Gln Ala Trp Val Gln Met Val Ala Arg Tyr Gly Lys Pro
660 665 670
Asn His Leu Asn Glu Arg Glu Gly Arg Gly Phe Ala Val Val Gly Tyr
675 680 685
Gly Lys Leu Gly Gly Trp Glu Leu Gly Tyr Ser Ser Asp Leu Asp Leu
690 695 700
Ile Phe Leu His Asp Cys Pro Met Asp Ala Met Thr Asp Gly Glu Arg
705 710 715 720
Glu Ile Asp Gly Arg Gln Phe Tyr Leu Arg Leu Ala Gln Arg Ile Met
725 730 735
His Leu Phe Ser Thr Arg Thr Ser Ser Gly Ile Leu Tyr Glu Val Asp
740 745 750
Ala Arg Leu Arg Pro Ser Gly Ala Ala Gly Met Leu Val Thr Ser Ala
755 760 765
Glu Ala Phe Ala Asp Tyr Gln Lys Asn Glu Ala Trp Thr Trp Glu His
770 775 780
Gln Ala Leu Val Arg Ala Arg Val Val Tyr Gly Asp Pro Gln Leu Thr
785 790 795 800
Ala His Phe Asp Ala Val Arg Arg Glu Ile Met Thr Leu Pro Arg Glu
805 810 815
Gly Lys Thr Leu Gln Thr Glu Val Arg Glu Met Arg Glu Lys Met Arg

```
            820                 825                 830
Ala His Leu Gly Asn Lys His Arg Asp Arg Phe Asp Ile Lys Ala Asp
        835                 840                 845
Glu Gly Gly Ile Thr Asp Ile Glu Phe Ile Thr Gln Tyr Leu Val Leu
    850                 855                 860
Arg Tyr Ala His Glu Lys Pro Lys Leu Thr Arg Trp Ser Asp Asn Val
865                 870                 875                 880
Arg Ile Leu Glu Leu Leu Ala Gln Asn Asp Ile Met Glu Glu Gln Glu
                885                 890                 895
Ala Met Ala Leu Thr Arg Ala Tyr Thr Thr Leu Arg Asp Glu Leu His
            900                 905                 910
His Leu Ala Leu Gln Glu Leu Pro Gly His Val Ser Glu Asp Cys Phe
        915                 920                 925
Thr Ala Glu Arg Glu Leu Val Arg Ala Ser Trp Gln Lys Trp Leu Val
    930                 935                 940
Glu Glu
945


<210> 471
<211> 433
<212> PRT
<213> Escherichia coli

<400> 471
Met Ala Gln Glu Ile Glu Leu Lys Phe Ile Val Asn His Ser Ala Val
 1               5                  10                  15
Glu Ala Leu Arg Asp His Leu Asn Thr Leu Gly Gly Glu His His Asp
            20                  25                  30
Pro Val Gln Leu Leu Asn Ile Tyr Tyr Glu Thr Pro Asp Asn Trp Leu
        35                  40                  45
Arg Gly His Asp Met Gly Leu Arg Ile Arg Gly Glu Asn Gly Arg Tyr
    50                  55                  60
Glu Met Thr Met Lys Val Ala Gly Arg Val Thr Gly Gly Leu His Gln
65                  70                  75                  80
Arg Pro Glu Tyr Asn Val Ala Leu Ser Glu Pro Thr Leu Asp Leu Ala
                85                  90                  95
Gln Leu Pro Thr Glu Val Trp Pro Asn Gly Glu Leu Pro Ala Asp Leu
            100                 105                 110
Ala Ser Arg Val Gln Pro Leu Phe Ser Thr Asp Phe Tyr Arg Glu Lys
        115                 120                 125
Trp Leu Val Ala Val Asp Gly Ser Gln Ile Glu Ile Ala Leu Asp Gln
    130                 135                 140
Gly Glu Val Lys Ala Gly Glu Phe Ala Glu Pro Ile Cys Glu Leu Glu
145                 150                 155                 160
Leu Glu Leu Leu Ser Gly Asp Thr Arg Ala Val Leu Lys Leu Ala Asn
                165                 170                 175
Gln Leu Val Ser Gln Thr Gly Leu Arg Gln Gly Ser Leu Ser Lys Ala
            180                 185                 190
Ala Arg Gly Tyr His Leu Ala Gln Gly Asn Pro Ala Arg Glu Ile Lys
        195                 200                 205
Pro Thr Thr Ile Leu His Val Ala Ala Lys Ala Asp Val Glu Gln Gly
    210                 215                 220
Leu Glu Ala Ala Leu Glu Leu Ala Leu Ala Gln Trp Gln Tyr His Glu
225                 230                 235                 240
Glu Leu Trp Val Arg Gly Asn Asp Ala Ala Lys Glu Gln Val Leu Ala
                245                 250                 255
Ala Ile Ser Leu Val Arg His Thr Leu Met Leu Phe Gly Gly Ile Val
            260                 265                 270
Pro Arg Lys Ala Ser Thr His Leu Arg Asp Leu Leu Thr Gln Cys Glu
        275                 280                 285
```

Ala Thr Ile Ala Ser Ala Val Ser Ala Val Thr Ala Val Tyr Ser Thr
290 295 300
Glu Thr Ala Met Ala Lys Leu Ala Leu Thr Glu Trp Leu Val Ser Lys
305 310 315 320
Ala Trp Gln Pro Phe Leu Asp Ala Lys Ala Gln Gly Lys Ile Ser Asp
325 330 335
Ser Phe Lys Arg Phe Ala Asp Ile His Leu Ser Arg His Ala Ala Glu
340 345 350
Leu Lys Ser Val Phe Cys Gln Pro Leu Gly Asp Arg Tyr Arg Asp Gln
355 360 365
Leu Pro Arg Leu Thr Arg Asp Ile Asp Ser Ile Leu Leu Leu Ala Gly
370 375 380
Tyr Tyr Asp Pro Val Val Ala Gln Ala Trp Leu Glu Asn Trp Gln Gly
385 390 395 400
Leu His His Ala Ile Ala Thr Gly Gln Arg Ile Glu Ile Glu His Phe
405 410 415
Arg Asn Glu Ala Asn Asn Gln Glu Pro Phe Trp Leu His Ser Gly Lys
420 425 430
Arg

<210> 472
<211> 207
<212> PRT
<213> Escherichia coli

<400> 472
Met Leu Asn Lys Leu Ser Leu Leu Leu Lys Asp Ala Gly Ile Ser Leu
1 5 10 15
Thr Asp His Gln Lys Asn Gln Leu Ile Ala Tyr Val Asn Met Leu His
20 25 30
Lys Trp Asn Lys Ala Tyr Asn Leu Thr Ser Val Arg Asp Pro Asn Glu
35 40 45
Met Leu Val Arg His Ile Leu Asp Ser Ile Val Val Ala Pro Tyr Leu
50 55 60
Gln Gly Glu Arg Phe Ile Asp Val Gly Thr Gly Pro Gly Leu Pro Gly
65 70 75 80
Ile Pro Leu Ser Ile Val Arg Pro Glu Ala His Phe Thr Leu Leu Asp
85 90 95
Ser Leu Gly Lys Arg Val Arg Phe Leu Arg Gln Val Gln His Glu Leu
100 105 110
Lys Leu Glu Asn Ile Glu Pro Val Gln Ser Arg Val Glu Glu Phe Pro
115 120 125
Ser Glu Pro Pro Phe Asp Gly Val Ile Ser Arg Ala Phe Ala Ser Leu
130 135 140
Asn Asp Met Val Ser Trp Cys His His Leu Pro Gly Glu Gln Gly Arg
145 150 155 160
Phe Tyr Ala Leu Lys Gly Gln Met Pro Glu Asp Glu Ile Ala Leu Leu
165 170 175
Pro Glu Glu Tyr Gln Val Glu Ser Val Val Lys Leu Gln Val Pro Ala
180 185 190
Leu Asp Gly Glu Arg His Leu Val Val Ile Lys Ala Asn Lys Ile
195 200 205

<210> 473
<211> 629
<212> PRT
<213> Escherichia coli

<400> 473
Met Phe Tyr Pro Asp Pro Phe Asp Val Ile Ile Ile Gly Gly Gly His
1 5 10 15
Ala Gly Thr Glu Ala Ala Met Ala Ala Ala Arg Met Gly Gln Gln Thr
20 25 30
Leu Leu Leu Thr His Asn Ile Asp Thr Leu Gly Gln Met Ser Cys Asn
35 40 45
Pro Ala Ile Gly Gly Ile Gly Lys Gly His Leu Val Lys Glu Val Asp
50 55 60
Ala Leu Gly Gly Leu Met Ala Lys Ala Ile Asp Gln Ala Gly Ile Gln
65 70 75 80
Phe Arg Ile Leu Asn Ala Ser Lys Gly Pro Ala Val Arg Ala Thr Arg
85 90 95
Ala Gln Ala Asp Arg Val Leu Tyr Arg Gln Ala Val Arg Thr Ala Leu
100 105 110
Glu Asn Gln Pro Asn Leu Met Ile Phe Gln Gln Ala Val Glu Asp Leu
115 120 125
Ile Val Glu Asn Asp Arg Val Val Gly Ala Val Thr Gln Met Gly Leu
130 135 140
Lys Phe Arg Ala Lys Ala Val Val Leu Thr Val Gly Thr Phe Leu Asp
145 150 155 160
Gly Lys Ile His Ile Gly Leu Asp Asn Tyr Ser Gly Gly Arg Ala Gly
165 170 175
Asp Pro Pro Ser Ile Pro Leu Ser Arg Arg Leu Arg Glu Leu Pro Leu
180 185 190
Arg Val Gly Arg Leu Lys Thr Gly Thr Pro Pro Arg Ile Asp Ala Arg
195 200 205
Thr Ile Asp Phe Ser Val Leu Ala Gln Gln His Gly Asp Asn Pro Met
210 215 220
Pro Val Phe Ser Phe Met Gly Asn Ala Ser Gln His Pro Gln Gln Val
225 230 235 240
Pro Cys Tyr Ile Thr His Thr Asn Glu Lys Thr His Asp Val Ile Arg
245 250 255
Ser Asn Leu Asp Arg Ser Pro Met Tyr Ala Gly Val Ile Glu Gly Val
260 265 270
Gly Pro Arg Tyr Cys Pro Ser Ile Glu Asp Lys Val Met Arg Phe Ala
275 280 285
Asp Arg Asn Gln His Gln Ile Phe Leu Glu Pro Glu Gly Leu Thr Ser
290 295 300
Asn Glu Ile Tyr Pro Asn Gly Ile Ser Thr Ser Leu Pro Phe Asp Val
305 310 315 320
Gln Met Gln Ile Val Arg Ser Met Gln Gly Met Glu Asn Ala Lys Ile
325 330 335
Val Arg Pro Gly Tyr Ala Ile Glu Tyr Asp Phe Phe Asp Pro Arg Asp
340 345 350
Leu Lys Pro Thr Leu Glu Ser Lys Phe Ile Gln Gly Leu Phe Phe Ala
355 360 365
Gly Gln Ile Asn Gly Thr Thr Gly Tyr Glu Glu Ala Ala Ala Gln Gly
370 375 380
Leu Leu Ala Gly Leu Asn Ala Ala Arg Leu Ser Ala Asp Lys Glu Gly
385 390 395 400
Trp Ala Pro Ala Arg Ser Gln Ala Tyr Leu Gly Val Leu Val Asp Asp
405 410 415
Leu Cys Thr Leu Gly Thr Lys Glu Pro Tyr Arg Met Phe Thr Ser Arg
420 425 430
Ala Glu Tyr Arg Leu Met Leu Arg Glu Asp Asn Ala Asp Leu Arg Leu
435 440 445
Thr Glu Ile Gly Arg Glu Leu Gly Leu Val Asp Asp Glu Arg Trp Ala
450 455 460
Arg Phe Asn Glu Lys Leu Glu Asn Ile Glu Arg Glu Arg Gln Arg Leu
465 470 475 480

Lys Ser Thr Trp Val Thr Pro Ser Ala Glu Ala Ala Ala Glu Val Asn
485 490 495
Ala His Leu Thr Ala Pro Leu Ser Arg Glu Ala Ser Gly Glu Asp Leu
500 505 510
Leu Arg Arg Pro Glu Met Thr Tyr Glu Lys Leu Thr Thr Leu Thr Pro
515 520 525
Phe Ala Pro Ala Leu Thr Asp Glu Gln Ala Ala Glu Gln Val Glu Ile
530 535 540
Gln Val Lys Tyr Glu Gly Tyr Ile Ala Arg Gln Gln Asp Glu Ile Glu
545 550 555 560
Lys Gln Leu Arg Asn Glu Asn Thr Leu Leu Pro Ala Thr Leu Asp Tyr
565 570 575
Arg Gln Val Ser Gly Leu Ser Asn Glu Val Ile Ala Lys Leu Asn Asp
580 585 590
His Lys Pro Ala Ser Ile Gly Gln Ala Ser Arg Ile Ser Gly Val Thr
595 600 605
Pro Ala Ala Ile Ser Ile Leu Leu Val Trp Leu Lys Lys Gln Gly Met
610 615 620
Leu Arg Arg Ser Ala
625

<210> 474
<211> 147
<212> PRT
<213> Escherichia coli

<400> 474
Met Ala Asp Ile Thr Leu Ile Ser Gly Ser Thr Leu Gly Gly Ala Glu
1 5 10 15
Tyr Val Ala Glu His Leu Ala Glu Lys Leu Glu Glu Ala Gly Phe Thr
20 25 30
Thr Glu Thr Leu His Gly Pro Leu Leu Glu Asp Leu Pro Ala Ser Gly
35 40 45
Ile Trp Leu Val Ile Ser Ser Thr His Gly Ala Gly Asp Ile Pro Asp
50 55 60
Asn Leu Ser Pro Phe Tyr Glu Ala Leu Gln Glu Gln Lys Pro Asp Leu
65 70 75 80
Ser Ala Val Arg Phe Gly Ala Ile Gly Ile Gly Ser Arg Glu Tyr Asp
85 90 95
Thr Phe Cys Gly Ala Ile Asp Lys Leu Glu Ala Glu Leu Lys Asn Ser
100 105 110
Gly Ala Lys Gln Thr Gly Glu Thr Leu Lys Ile Asn Ile Leu Asp His
115 120 125
Asp Ile Pro Glu Asp Pro Ala Glu Glu Trp Leu Gly Ser Trp Val Asn
130 135 140
Leu Leu Lys
145

<210> 475
<211> 151
<212> PRT
<213> Escherichia coli

<400> 475
Met Gly Gln Arg Asn Val Ser Leu Met Glu Lys Lys Met Lys Lys Gly
1 5 10 15
Thr Val Leu Asn Ser Asp Ile Ser Ser Val Ile Ser Arg Leu Gly His
20 25 30
Thr Asp Thr Leu Val Val Cys Asp Ala Gly Leu Pro Ile Pro Lys Ser

35 40 45
Thr Thr Arg Ile Asp Met Ala Leu Thr Gln Gly Val Pro Ser Phe Met
50 55 60
Gln Val Leu Gly Val Val Thr Asn Glu Met Gln Val Glu Ala Ala Ile
65 70 75 80
Ile Ala Glu Glu Ile Lys His His Asn Pro Gln Leu His Glu Thr Leu
85 90 95
Leu Thr His Leu Glu Gln Leu Gln Lys His Gln Gly Asn Thr Ile Glu
100 105 110
Ile Arg Tyr Thr Thr His Glu Gln Phe Lys Gln Gln Thr Ala Glu Ser
115 120 125
Gln Ala Val Ile Arg Ser Gly Glu Cys Ser Pro Tyr Ala Asn Ile Ile
130 135 140
Leu Cys Ala Gly Val Thr Phe
145 150

<210> 476
<211> 501
<212> PRT
<213> Escherichia coli

<400> 476
Met Glu Ala Leu Leu Gln Leu Lys Gly Ile Asp Lys Ala Phe Pro Gly
1 5 10 15
Val Lys Ala Leu Ser Gly Ala Ala Leu Asn Val Tyr Pro Gly Arg Val
20 25 30
Met Ala Leu Val Gly Glu Asn Gly Ala Gly Lys Ser Thr Met Met Lys
35 40 45
Val Leu Thr Gly Ile Tyr Thr Arg Asp Ala Gly Thr Leu Leu Trp Leu
50 55 60
Gly Lys Glu Thr Thr Phe Thr Gly Pro Lys Ser Ser Gln Glu Ala Gly
65 70 75 80
Ile Gly Ile Ile His Gln Glu Leu Asn Leu Ile Pro Gln Leu Thr Ile
85 90 95
Ala Glu Asn Ile Phe Leu Gly Arg Glu Phe Val Asn Arg Phe Gly Lys
100 105 110
Ile Asp Trp Lys Thr Met Tyr Ala Glu Ala Asp Lys Leu Leu Ala Lys
115 120 125
Leu Asn Leu Arg Phe Lys Ser Asp Lys Leu Val Gly Asp Leu Ser Ile
130 135 140
Gly Asp Gln Gln Met Val Glu Ile Ala Lys Val Leu Ser Phe Glu Ser
145 150 155 160
Lys Val Ile Ile Met Asp Glu Pro Thr Asp Ala Leu Thr Asp Thr Glu
165 170 175
Thr Glu Ser Leu Phe Arg Val Ile Arg Glu Leu Lys Ser Gln Gly Arg
180 185 190
Gly Ile Val Tyr Ile Ser His Arg Met Lys Glu Ile Phe Glu Ile Cys
195 200 205
Asp Asp Val Thr Val Phe Arg Asp Gly Gln Phe Ile Ala Glu Arg Glu
210 215 220
Val Ala Ser Leu Thr Glu Asp Ser Leu Ile Glu Met Met Val Gly Arg
225 230 235 240
Lys Leu Glu Asp Gln Tyr Pro His Leu Asp Lys Ala Pro Gly Asp Ile
245 250 255
Arg Leu Lys Val Asp Asn Leu Cys Gly Pro Gly Val Asn Asp Val Ser
260 265 270
Phe Thr Leu Arg Lys Gly Glu Ile Leu Gly Val Ser Gly Leu Met Gly
275 280 285
Ala Gly Arg Thr Glu Leu Met Lys Val Leu Tyr Gly Ala Leu Pro Arg
290 295 300

Thr Ser Gly Tyr Val Thr Leu Asp Gly His Glu Val Val Thr Arg Ser
305 310 315 320
Pro Gln Asp Gly Leu Ala Asn Gly Ile Val Tyr Ile Ser Glu Asp Arg
325 330 335
Lys Arg Asp Gly Leu Val Leu Gly Met Ser Val Lys Glu Asn Met Ser
340 345 350
Leu Thr Ala Leu Arg Tyr Phe Ser Arg Ala Gly Gly Ser Leu Lys His
355 360 365
Ala Asp Glu Gln Gln Ala Val Ser Asp Phe Ile Arg Leu Phe Asn Val
370 375 380
Lys Thr Pro Ser Met Glu Gln Ala Ile Gly Leu Leu Ser Gly Gly Asn
385 390 395 400
Gln Gln Lys Val Ala Ile Ala Arg Gly Leu Met Thr Arg Pro Lys Val
405 410 415
Leu Ile Leu Asp Glu Pro Thr Arg Gly Val Asp Val Gly Ala Lys Lys
420 425 430
Glu Ile Tyr Gln Leu Ile Asn Gln Phe Lys Ala Asp Gly Leu Ser Ile
435 440 445
Ile Leu Val Ser Ser Glu Met Pro Glu Val Leu Gly Met Ser Asp Arg
450 455 460
Ile Ile Val Met His Glu Gly His Leu Ser Gly Glu Phe Thr Arg Glu
465 470 475 480
Gln Ala Thr Gln Glu Val Leu Met Ala Ala Ala Val Gly Lys Leu Asn
485 490 495
Arg Val Asn Gln Glu
500

<210> 477
<211> 321
<212> PRT
<213> Escherichia coli

<400> 477
Met Thr Thr Gln Thr Val Ser Gly Arg Arg Tyr Phe Thr Lys Ala Trp
1 5 10 15
Leu Met Glu Gln Lys Ser Leu Ile Ala Leu Leu Val Leu Ile Ala Ile
20 25 30
Val Ser Thr Leu Ser Pro Asn Phe Phe Thr Ile Asn Asn Leu Phe Asn
35 40 45
Ile Leu Gln Gln Thr Ser Val Asn Ala Ile Met Ala Val Gly Met Thr
50 55 60
Leu Val Ile Leu Thr Ser Gly Ile Asp Leu Ser Val Gly Ser Leu Leu
65 70 75 80
Ala Leu Thr Gly Ala Val Ala Ala Ser Ile Val Gly Ile Glu Val Asn
85 90 95
Ala Leu Val Ala Val Ala Ala Ala Leu Ala Leu Gly Ala Ala Ile Gly
100 105 110
Ala Val Thr Gly Val Ile Val Ala Lys Gly Arg Val Gln Ala Phe Ile
115 120 125
Ala Thr Leu Val Met Met Leu Leu Leu Arg Gly Val Thr Met Val Tyr
130 135 140
Thr Asn Gly Ser Pro Val Asn Thr Gly Phe Thr Glu Asn Ala Asp Leu
145 150 155 160
Phe Gly Trp Phe Gly Ile Gly Arg Pro Leu Gly Val Pro Thr Pro Val
165 170 175
Trp Ile Met Gly Ile Val Phe Leu Ala Ala Trp Tyr Met Leu His His
180 185 190
Thr Arg Leu Gly Arg Tyr Ile Tyr Ala Leu Gly Gly Asn Glu Ala Ala
195 200 205
Thr Arg Leu Ser Gly Ile Asn Val Asn Lys Ile Lys Ile Ile Val Tyr

```
    210                 215                 220
Ser Leu Cys Gly Leu Leu Ala Ser Leu Ala Gly Ile Ile Glu Val Ala
225                 230                 235                 240
Arg Leu Ser Ser Ala Gln Pro Thr Ala Gly Thr Gly Tyr Glu Leu Asp
                245                 250                 255
Ala Ile Ala Ala Val Val Leu Gly Gly Thr Ser Leu Ala Gly Gly Lys
            260                 265                 270
Gly Arg Ile Val Gly Thr Leu Ile Gly Ala Leu Ile Leu Gly Phe Leu
        275                 280                 285
Asn Asn Gly Leu Asn Leu Leu Gly Val Ser Ser Tyr Tyr Gln Met Ile
    290                 295                 300
Val Lys Ala Val Val Ile Leu Leu Ala Val Leu Val Asp Asn Lys Lys
305                 310                 315                 320
Gln
```

<210> 478
<211> 296
<212> PRT
<213> Escherichia coli

<400> 478

```
Met Asn Met Lys Lys Leu Ala Thr Leu Val Ser Ala Val Ala Leu Ser
 1               5                  10                  15
Ala Thr Val Ser Ala Asn Ala Met Ala Lys Asp Thr Ile Ala Leu Val
            20                  25                  30
Val Ser Thr Leu Asn Asn Pro Phe Phe Val Ser Leu Lys Asp Gly Ala
        35                  40                  45
Gln Lys Glu Ala Asp Lys Leu Gly Tyr Asn Leu Val Val Leu Asp Ser
    50                  55                  60
Gln Asn Asn Pro Ala Lys Glu Leu Ala Asn Val Gln Asp Leu Thr Val
65                  70                  75                  80
Arg Gly Thr Lys Ile Leu Leu Ile Asn Pro Thr Asp Ser Asp Ala Val
                85                  90                  95
Gly Asn Ala Val Lys Met Ala Asn Gln Ala Asn Ile Pro Val Ile Thr
            100                 105                 110
Leu Asp Arg Gln Ala Thr Lys Gly Glu Val Val Ser His Ile Ala Ser
        115                 120                 125
Asp Asn Val Leu Gly Gly Lys Ile Ala Gly Asp Tyr Ile Ala Lys Lys
    130                 135                 140
Ala Gly Glu Gly Ala Lys Val Ile Glu Leu Gln Gly Ile Ala Gly Thr
145                 150                 155                 160
Ser Ala Ala Arg Glu Arg Gly Glu Gly Phe Gln Gln Ala Val Ala Ala
                165                 170                 175
His Lys Phe Asn Val Leu Ala Ser Gln Pro Ala Asp Phe Asp Arg Ile
            180                 185                 190
Lys Gly Leu Asn Val Met Gln Asn Leu Leu Thr Ala His Pro Asp Val
        195                 200                 205
Gln Ala Val Phe Ala Gln Asn Asp Glu Met Ala Leu Gly Ala Leu Arg
    210                 215                 220
Ala Leu Gln Thr Ala Gly Lys Ser Asp Val Met Val Val Gly Phe Asp
225                 230                 235                 240
Gly Thr Pro Asp Gly Glu Lys Ala Val Asn Asp Gly Lys Leu Ala Ala
                245                 250                 255
Thr Ile Ala Gln Leu Pro Asp Gln Ile Gly Ala Lys Gly Val Glu Thr
            260                 265                 270
Ala Asp Lys Val Leu Lys Gly Glu Lys Val Gln Ala Lys Tyr Pro Val
        275                 280                 285
Asp Leu Lys Leu Val Val Lys Gln
    290                 295
```

<210> 479
<211> 309
<212> PRT
<213> Escherichia coli

<400> 479
Met Gln Asn Ala Gly Ser Leu Val Val Leu Gly Ser Ile Asn Ala Asp
1 5 10 15
His Ile Leu Asn Leu Gln Ser Phe Pro Thr Pro Gly Glu Thr Val Thr
20 25 30
Gly Asn His Tyr Gln Val Ala Phe Gly Gly Lys Gly Ala Asn Gln Ala
35 40 45
Val Ala Ala Gly Arg Ser Gly Ala Asn Ile Ala Phe Ile Ala Cys Thr
50 55 60
Gly Asp Asp Ser Ile Gly Glu Ser Val Arg Gln Gln Leu Ala Thr Asp
65 70 75 80
Asn Ile Asp Ile Thr Pro Val Ser Val Ile Lys Gly Glu Ser Thr Gly
85 90 95
Val Ala Leu Ile Phe Val Asn Gly Glu Gly Glu Asn Val Ile Gly Ile
100 105 110
His Ala Gly Ala Asn Ala Ala Leu Ser Pro Ala Leu Val Glu Ala Gln
115 120 125
Arg Glu Arg Ile Ala Asn Ala Ser Ala Leu Leu Met Gln Leu Glu Ser
130 135 140
Pro Leu Glu Ser Val Met Ala Ala Ala Lys Ile Ala His Gln Asn Lys
145 150 155 160
Thr Ile Val Ala Leu Asn Pro Ala Pro Ala Arg Glu Leu Pro Asp Glu
165 170 175
Leu Leu Ala Leu Val Asp Ile Ile Thr Pro Asn Glu Thr Glu Ala Glu
180 185 190
Lys Leu Thr Gly Ile Arg Val Glu Asn Asp Glu Asp Ala Ala Lys Ala
195 200 205
Ala Gln Val Leu His Glu Lys Gly Ile Arg Thr Val Leu Ile Thr Leu
210 215 220
Gly Ser Arg Gly Val Trp Ala Ser Val Asn Gly Glu Gly Gln Arg Val
225 230 235 240
Pro Gly Phe Arg Val Gln Ala Val Asp Thr Ile Ala Ala Gly Asp Thr
245 250 255
Phe Asn Gly Ala Leu Ile Thr Ala Leu Leu Glu Glu Lys Pro Leu Pro
260 265 270
Glu Ala Ile Arg Phe Ala His Ala Ala Ala Ala Ile Ala Val Thr Arg
275 280 285
Lys Gly Ala Gln Pro Ser Val Pro Trp Arg Glu Glu Ile Asp Ala Phe
290 295 300
Leu Asp Arg Gln Arg
305

<210> 480
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<223> Linker containing T7 terminator

<400> 480
gtctagcata accccttggg gcctctaaac gggtccttga ggggtttttt ga 52

<210> 481
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> Linker containing T7 terminator

<400> 481
agcttcaaaa aacccctcaa ggacccgttt agaggcccca aggggttatg ctagactgca 60

U.S.S.N.: 10/521,300
Filed: 1/14/05
Case No.: 21126YP

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 July 2001 (05.07.2001)

PCT

(10) International Publication Number
WO 01/48209 A3

(51) International Patent Classification[7]: **C12N 15/31**, 15/11, 15/10, C07K 14/245

(21) International Application Number: PCT/US00/34419

(22) International Filing Date: 19 December 2000 (19.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/173,005 23 December 1999 (23.12.1999) US

(71) Applicant *(for all designated States except US)*: **ELITRA PHARMACEUTICALS, INC.** [US/US]; Suite A, 3510 Dunhill Street, San Diego, CA 92121 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **FORSYTH, R., Allyn** [US/US]; 1135 Beryl Street, San Diego, CA 92109 (US). **OHLSEN, Kari, L.** [US/US]; 3560 Vista De La Orilla, San Diego, CA 92117 (US). **ZYSKIND, Judith, W.** [US/US]; 8415 La Jolla Scenic Drive, La Jolla, CA 92037 (US).

(74) Agent: ISRAELSEN, Ned, A.: 16th Floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *with amended claims*

(88) Date of publication of the international search report: 2 May 2002

Date of publication of the amended claims: 6 June 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/48209 A3

(54) Title: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION OF E. COLI

(57) Abstract: The sequences of nucleic acids encoding proteins required for E. coli proliferation are disclosed. The nucleic acids can also be used to screen for homologous genes that are required for proliferation in microorganisms other than E. coli. The nucleic acids can also be used to design expression vectors and secretion vectors. The nucleic acids can be used to express proteins or portions thereof, to obtain antibodies capable of specifically binding to the expressed proteins, and to use those expressed proteins as a screen to isolate candidate molecules for rational drug discovery programs. The nucleic acids of the present invention can also be used in various assay systems to screen for antimicrobial agents.

AMENDED CLAIMS

[received by the International Bureau on 14 January 2002 (14.01.02);
original claims 1-131 replaced by amended claims 1-131 (14 pages)]

1. A purified or isolated nucleic acid sequence consisting essentially of one the sequence of nucleotides of SEQ ID NOs: 1-93, wherein expression of said nucleic acid in a microorganism is capable of inhibiting proliferation of a microorganism.

2. The nucleic acid sequence of Claim 1, wherein said nucleic acid sequence is complementary to at least a portion of the nucleotide sequence of the coding strand of a gene whose expression is required for proliferation of a microorganism.

3. The nucleic acid of Claim 1, wherein said nucleic acid sequence has a nucleotide seuqence complementary to at least a portion of the nucleotide sequence of an RNA required for proliferation of a microorganism.

4. The nucleic acid of Claim 3, wherein the nucleotide sequence of said RNA encodes more than one gene product.

5. A purified or isolated nucleic acid comprising a fragment of one of the nucleotide seuqence of SEQ ID NOs.: 1-93, said fragment selected from the group consisting of fragments comprising at least 10, at least 20, at least 25, at least 30, at least 50 and more than 50 consecutive nucleotides of one of the nucleotide sequences of SEQ ID NOs: 1-93.

6. A vector comprising a promoter operably linked to the nucleic acid sequence of Claims 1,2,3,4, or 5.

7. The vector of Claim 6, wherein said promoter is active in a microorganism selected from the group consisting of Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis and any species falling within the genera of any of the above species.

8. A host cell containing the vector of Claim 6.

9. A purified or isolated nucleic acid consisting essentially of the coding sequence of one of SEQ ID NOs: 105-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286.

10. A fragment of the nucleic acid of Claim 8, said fragment comprising at least 10, at least 20, at least 25, at least 30, at least 50 or more than 50 consecutive nucleotides of one of SEQ ID NOs: 105-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, and 279-286.

11. A vector comprising a promoter operably linked to the nucleic acid of Claim 9 or Claim 10.

12. A purified or isolated antisense nucleic acid comprising a nucleic acid sequence complementary to at least a portion of an intragenic sequence, intergenic sequence, sequences spanning at least a portion of two or more genes, 5' noncoding region, or 3' noncoding region within an operon comprising a proliferation-required gene whose activity or expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93.

13. A purified or isolated nucleic acid comprising a nucleic acid having at least 70% identity to a sequence selected from the group consisting of SEQ ID NOs.: 1-93, fragments comprising at least 25 consecutive nucleotides of SEQ ID NOs.: 1-93, the sequences complementary to SEQ ID NOs.: 1-93 and the sequences complementary to fragments comprising at least 25 consecutive nucleotides of SEQ ID NOs.: 1-93 as determined using BLASTN version 2.0 with the default parameters.

14. The nucleic acid of Claim 13, wherein said nucleic acid is from an organism selected from the group consisting of Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis and any species falling within the genera of any of the above species.

15. A vector comprising a promoter operably linked to a nucleic acid encoding a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93.

16. A host cell containing the vector of Claim 15.

17. The vector of Claim 15, wherein said polypeptide comprises a polypeptide

comprising a sequence selected from the group consisting of SEQ ID NOs: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

18. A purified or isolated polypeptide comprising a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93, or a fragment selected from the group consisting of fragments comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of one of the said polypeptides.

19. The polypeptide of Claim 18, wherein said polypeptide comprises a polypeptide comprising one of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 or a fragment comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

20. A purified or isolated polypeptide comprising a polypeptide having at least 25% identity to a polypeptide whose expression is inhibited by a sequence selected from the group consisting of SEQ ID NOs.: 1-93, or at least 25% identity to a fragment comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide whose expression is inhibited by a nucleic acid selected from the group consisting of SEQ ID NOs.: 1-93 as determined using FASTA version 3.0t78 with the default parameters.

21. The polypeptide of Claim 20, wherein said polypeptide has at least 25% identity to a polypeptide comprising one of SEQ ID NOs: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 or at least 25% identity to a fragment comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide comprising one of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479 as determined using FASTA version 3.0t78 with the default parameters.

22. An antibody capable of specifically binding the polypeptide of one of Claims 18-21.

23. A method of producing a polypeptide, comprising introducing a vector comprising a promoter operably linked to a nucleic acid encoding a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93 into a cell and expressing said polypeptide.

24. The method of Claim 23, further comprising the step of isolating said polypeptide.

25. The method of Claim 23, wherein said polypeptide comprises a sequence selected from the group consisting of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

26. A method of inhibiting proliferation of a microorganism comprising inhibiting the

activity or reducing the amount of a gene product whose expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 or inhibiting the activity or reducing the amount of a nucleic acid encoding said gene product.

27. The method of Claim 26, wherein said gene product comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

28. A method for identifying a compound which influences the activity of a gene product required for proliferation, said gene product comprising a gene product whose expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising:

contacting said gene product with a candidate compound; and

determining whether said compound influences the activity of said gene product.

29. The method of Claim 28, wherein said gene product is a polypeptide and said activity is an enzymatic activity.

30. The method of Claim 28, wherein said gene product is a polypeptide and said activity is a carbon compound catabolism activity.

31. The method of Claim 28, wherein said gene product is a polypeptide and said activity is a biosynthetic activity.

32. The method of Claim 28, wherein said gene product is a polypeptide and said activity is a transporter activity.

33. The method of Claim 28, wherein said gene product is a polypeptide and said activity is a transcriptional activity.

34. The method of Claim 28, wherein said gene product is a polypeptide and said activity is a DNA replication activity.

35. The method of Claim 28, wherein said gene product is a polypeptide and said activity is a cell division activity.

36. A compound identified using the method of Claim 28.

37. The method of Claim 28, wherein said gene product is a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

38. A method for identifying a compound or nucleic acid having the ability to reduce the activity or level of a gene product required for proliferation, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising:

(a) providing a target that is a gene or RNA, wherein said target comprises a nucleic acid encoding said gene product;

(b) contacting said target with a candidate compound or nucleic acid; and

(c) measuring an activity of said target.

39. The method of Claim 38, wherein said target is a messenger RNA molecule and said activity is translation of said messenger RNA.

40. The method of Claim 38, wherein said target is a messenger RNA molecule and said activity is transcription of a gene encoding said messenger RNA.

41. The method of Claim 38, wherein said target is a gene and said activity is transcription of said gene.

42. The method of Claim 38, wherein said target is a nontranslated RNA and said activity is processing or folding of said nontranslated RNA or assembly of said nontranslated RNA into a protein/RNA complex.

43. The method of Claim 38. wherein said target gene or RNA encodes a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466. 468 and 472-479.

44. A compound or nucleic acid identified using the method of Claim 38.

45. A method for identifying a compound which reduces the activity or level of a gene product required for proliferation of a microorganism, wherein the activity or expression of said gene product is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93. said method comprising the steps of:

(a) expressing a sub-lethal level of an antisense nucleic acid complementary to a nucleic acid encoding said gene product in a cell to reduce the activity or amount of said gene product in said cell, thereby producing a sensitized cell;

(b) contacting said sensitized cell with a compound; and

(c) determining whether said compound inhibits the growth of said sensitized cell.

46. The method of Claim 45. wherein said determining step comprises determining whether said compound inhibits the growth of said sensitized cell to a greater extent than said compound inhibits the growth of a nonsensitized cell.

47. The method of Claim 45. wherein said cell is selected from the group consisting of bacterial cells, fungal cells, plant cells. and animal cells.

48. The method of Claim 45, wherein said cell is a Gram negative bacterium.

49. The method of Claim 45. wherein said cell is an *E. coli* cell.

50. The method of Claim 45, wherein said cell is from an organism selected from the

group consisting of Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis. Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis and any species falling within the genera of any of the above species.

51. The method of Claim 45, wherein said antisense nucleic acid is transcribed from an inducible promoter.

52. The method of Claim 51, further comprising the step of contacting said cell with a concentration of inducer which induces said antisense nucleic acid to a sub-lethal level.

53. The method of Claim 45, wherein growth inhibition is measured by monitoring optical density of a culture growth solution.

54. The method of Claim 45, wherein said gene product is a polypeptide.

55. The method of Claim 54, wherein said polypeptide comprises a sequence selected from the group consisting of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

56. The method of Claim 45, wherein said gene product is an RNA.

57. A compound identified using the method of Claim 45.

58. A method for inhibiting cellular proliferation comprising introducing a compound with activity against a gene whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 or a compound with activity against the product of said gene into a population of cells expressing said gene.

59. The method of Claim 58, wherein said compound is an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, or a proliferation-inhibiting portion thereof.

60. The method of Claim 59, wherein said proliferation inhibiting portion of one of SEQ ID NOs.: 1-93 is a fragment comprising at least 10, at least 20, at least 25, at least 30, at least 50 or more than 51 consecutive nucleotides of one of SEQ ID NOs.: 1-93.

61. The method of Claim 58, wherein said population is a population selected from the group consisting of bacterial cells, fungal cells, plant cells, and animal cells.

62. The method of Claim 58, wherein said population is a population of Gram negative bacteria.

63. The method of Claim 58, wherein said population is a population of *E. coli* cells.

64. The method of Claim 58, wherein said population is a population selected from the group consisting of Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae,Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis and any species falling within the genera of any of the above species.

65. The method of Claim 58, wherein said gene encodes a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

66. A preparation comprising an effective concentration of an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, or a proliferation-inhibiting portion thereof in a pharmaceutically acceptable carrier.

67. The preparation of Claim 66, wherein said proliferation-inhibiting portion of one of SEQ ID NOs.: 1-93 comprises at least 10, at least 20, at least 25, at least 30, at least 50 or more than 50 consecutive nucleotides of one of SEQ ID NOs.: 1-93.

68. A method for inhibiting the activity or expression of a gene in an operon required for proliferation wherein the activity or expression of at least one gene in said operon is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising contacting a cell in a cell population with an antisense nucleic acid comprising at least a proliferation-inhibiting portion of said operon.

69. The method of Claim 68, wherein said antisense nucleic acid comprises a sequence selected from the group consisting of SEQ ID NOs.: 1-93 or a proliferation inhibiting portion thereof.

70. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a plasmid which expresses said antisense nucleic acid into said cell population.

71. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a phage which expresses said antisense nucleic acid into said cell population.

72. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by expressing said antisense nucleic acid from the chromosome of cells in said cell population.

73. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a promoter adjacent to a chromosomal copy of said antisense nucleic acid such that said promoter directs the synthesis of said antisense nucleic acid.

74. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a retron which expresses said antisense nucleic acid into said cell population.

75. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a ribozyme into said cell-population, wherein a binding portion of said ribozyme is complementary to said antisense oligonucleotide.

76. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by introducing a liposome comprising said antisense oligonucleotide into said cell.

77. The method of Claim 68, wherein said cell is contacted with said antisense nucleic acid by electroporation of said antisense nucleic acid.

78. The method of Claim 68, wherein said antisense nucleic acid is a fragment comprising at least 10, at least 20, at least 25, at least 30, at least 50 or more than 50 consecutive nucleotides of one of SEQ ID NOs.: 1-93.

79. The method of Claim 68 wherein said antisense nucleic acid is an oligonucleotide.

80. A method for identifying a gene which is required for proliferation of a microorganism comprising:

(a) contacting a microorganism other than *E. coli* with a nucleic acid selected from the group consisting of SEQ ID NOs.: 1-93;

(b) determining whether said nucleic acid inhibits proliferation of said microorganism; and

(c) identifying the gene in said microorganism which is inhibited by said nucleic acid.

81. The method of Claim 80, wherein said microorganism is a Gram negative bacterium.

82. The method of Claim 80 wherein said microorganism is selected from the group consisting of Aspergillus fumigatus, Bacillus anthracis, Burkholderia cepacia, Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata), Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida

pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae, Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes. Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi. Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei. Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum, Yersinia pestis and any species falling within the genera of any of the above species.

83. The method of Claim 80. further comprising introducing said nucleic acid into a vector functional in said microorganism prior to introducing said inhibitory nucleic acid into said microorganism.

84. A method for identifying a compound having the ability to inhibit proliferation of a microorganism comprising:

(a) identifying in a first microorganism a homolog of a gene or gene product present in a second microorganism which is different than said first microorganism, wherein the activity or level of said gene or gene product is inhibited by a nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs. 1-93 ;

(b) identifying an inhibitory nucleic acid sequence which inhibits the activity of said homolog in said first microorganism;

(c) contacting said first microorganism with a sub-lethal level of said inhibitory nucleic acid, thus sensitizing said first microorganism:

(d) contacting the sensitized microorganism of step (c) with a compound; and

(e) determining whether said compound inhibits proliferation of said sensitized microorganism.

85. The method of Claim 84. wherein said determining step comprises determining whether said compound inhibits proliferation of said sensitized microorganism to a greater extent than said compound inhibits proliferation of a nonsensitized microorganism.

86. The method of Claim 84 wherein step (a) comprises identifying a homologous nucleic acid to a gene or gene product whose activity or level is inhibited by a nucleic acid selected from the group consisting of SEQ ID NOs. 1-93 or a nucleic acid encoding a homologous polypeptide to a polypeptide whose activity or level is inhibited by a nucleic acid selected from the group consisting of SEQ ID NOs. 1-93 by using an algorithm selected from the group consisting of BLASTN version 2.0 with the default parameters and FASTA version 3.0t78 algorithm with the

default parameters to identify said homologous nucleic acid or said nucleic acid encoding a homologous polypeptide in a database.

87. The method of Claim 84 wherein said step (a) comprises identifying a homologous nucleic acid or a nucleic acid encoding a homologous polypeptide by identifying nucleic acids which hybridize to said first gene.

88. The method of Claim 84 wherein the step (a) comprises expressing a nucleic acid selected from the group consisting of SEQ ID NOs. 1-93 in said microorganism.

89. The method of Claim 84, wherein said inhibitory nucleic acid is an antisense nucleic acid.

90. The method of Claim 84, wherein said inhibitory nucleic acid comprises an antisense nucleic acid to a portion of said homolog.

91. The method of Claim 84, wherein said inhibitory nucleic acid comprises an antisense nucleic acid to a portion of the operon encoding said homolog.

92. The method of Claim 84, wherein the step of contacting the first microorganism with a sub-lethal level of said inhibitory nucleic acid comprises directly contacting said microorganism with said inhibitory nucleic acid.

93. The method of Claim 84, wherein the step of contacting the first microorganism with a sub-lethal level of said inhibitory nucleic acid comprises expressing an antisense nucleic acid to said homolog in said microorganism.

94. The method of Claim 84, wherein said gene product comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

95. A compound identified using the method of Claim 84.

96. A method of identifying a compound having the ability to inhibit proliferation comprising:

(a) contacting a microorganism other than *E. coli* with a sub-lethal level of a nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs. 1-93 or a portion thereof which inhibits the proliferation of *E. coli*, thus sensitizing said microorganism;

(b) contacting the sensitized microorganism of step (a) with a compound; and

(c) determining whether said compound inhibits proliferation of said sensitized microorganism.

97. The method of Claim 96, wherein said determining step comprises determining whether said compound inhibits proliferation of said sensitized microorganism to a greater extent than said compound inhibits proliferation of a nonsensitized microorganism.

AMENDED SHEET (ARTICLE 19)

98. A compound identified using the method of Claim 96.

99. A method for identifying a compound having activity against a biological pathway required for proliferation comprising:

(a) sensitizing a cell by expressing a sub-lethal level of an antisense nucleic acid complementary to a nucleic acid encoding a gene product required for proliferation, wherein the activity or expression of said gene product is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, in said cell to reduce the activity or amount of said gene product;

(b) contacting the sensitized cell with a compound; and

(c) determining whether said compound inhibits the growth of said sensitized cell.

100. The method of Claim 99. wherein said determining step comprises determining whether said compound inhibits the growth of said sensitized cell to a greater extent than said compound inhibits the growth of a nonsensitized cell.

101. The method of Claim 99. wherein said cell is selected from the group consisting of bacterial cells, fungal cells. plant cells. and animal cells.

102. The method of Claim 99, wherein said cell is a Gram negative bacterium.

103. The method of Claim 99, wherein said Gram negative bacterium is *E. coli.*

104. The method of Claim 99. wherein said cell is selected from the group consisting of Aspergillus fumigatus. Bacillus anthracis. Burkholderia cepacia. Campylobacter jejuni, Candida albicans, Candida glabrata (also called Torulopsis glabrata). Candida tropicalis, Candida parapsilosis, Candida guilliermondii, Candida krusei, Candida kefyr (also called Candida pseudotropicalis), Candida dubliniensis, Chlamydia pneumoniae, Chlamydia trachomatus, Clostridium botulinum, Clostridium difficile, Cryptococcus neoformans, Enterobacter cloacae, Enterococcus faecalis, Escherichia coli, Haemophilus influenzae. Helicobacter pylori, Klebsiella pneumoniae, Listeria monocytogenes, Mycobacterium leprae, Mycobacterium tuberculosis, Neisseria gonorrhoeae, Pseudomonas aeruginosa, Salmonella cholerasuis, Salmonella enterica, Salmonella paratyphi, Salmonella typhi, Salmonella typhimurium, Staphylococcus aureus, Klebsiella pneumoniae, Listeria monocytogenes, Moxarella catarrhalis. Shigella boydii. Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Pseudomonas aeruginosa, Staphylococcus epidermidis, Streptococcus pneumoniae, Treponema pallidum. Yersinia pestis and any species falling within the genera of any of the above species.

105. The method of Claim 99, wherein said antisense nucleic acid is transcribed from an inducible promoter.

106. The method of Claim 99, further comprising contacting the cell with an agent which

induces expression of said antisense nucleic acid from said inducible promoter, wherein said antisense nucleic acid is expressed at a sub-lethal level.

107. The method of Claim 99, wherein inhibition of proliferation is measured by monitoring the optical density of a liquid culture.

108. The method of Claim 99, wherein said gene product comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479

109. A compound identified using the method of Claim 99.

110. A method for identifying a compound having the ability to inhibit cellular proliferation comprising:

(a) contacting a cell with an agent which reduces the activity or level of a gene product required for proliferation of said cell, wherein said gene product is a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93;

(b) contacting said cell with a compound; and

(c) determining whether said compound reduces proliferation of said contacted cell.

111. The method of Claim 110, wherein said determining step comprises determining whether said compound reduces proliferation of said contacted cell to a greater extent than said compound reduces proliferation of cells which have not been contacted with said agent.

112. The method of Claim 110, wherein said agent which reduces the activity or level of a gene product required for proliferation of said cell comprises an antisense nucleic acid to a gene or operon required for proliferation.

113. The method of Claim 110, wherein said agent which reduces the activity or level of a gene product required for proliferation of said cell comprises a compound known to inhibit growth or proliferation of a microorganism.

114. The method of Claim 110, wherein said cell contains a mutation which reduces the activity or level of said gene product required for proliferation of said cell.

115. The method of Claim 114, wherein said mutation is a temperature sensitive mutation.

116. The method of Claim 110, wherein said gene product comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479

117. A compound identified using the method of Claim 110.

118. A method for identifying the biological pathway in which a proliferation-required gene or its gene product lies, wherein said gene or gene product comprises a gene or gene product

whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93, said method comprising:

(a) expressing a sub-lethal level of an antisense nucleic acid which inhibits the activity of said proliferation-required gene or gene product in a cell;

(b) contacting said cell with a compound known to inhibit growth or proliferation of a microorganism, wherein the biological pathway on which said compound acts is known; and

(c) determining whether said cell is sensitive to said compound.

119. The method of Claim 118, wherein said determining step comprises determining whether said cell has a substantially greater sensitivity to said compound than a cell which does not express said sub-lethal level of said antisense nucleic acid and wherein said gene or gene product lies in the same pathway on which said compound acts if said cell expressing said sub-lethal level of said antisense nucleic acid has a substantially greater sensitivity to said compound than said cell which does not express said sub-lethal level of said antisense nucleic acid.

120. The method of Claim 118, wherein said gene product comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479

121. A method for determining the biological pathway on which a test compound acts comprising:

(a) expressing a sub-lethal level of an antisense nucleic acid complementary to a proliferation-required nucleic acid in a cell, wherein the activity or expression of said proliferation-required nucleic acid is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 and wherein the biological pathway in which said proliferation-required nucleic acid or a protein encoded by said proliferation-required polypeptide lies is known,

(b) contacting said cell with said test compound; and

(c) determining whether said cell is sensitive to said test compound.

122. The method of Claim 121, wherein said determining step comprises determining whether said cell has a substantially greater sensitivity to said test compound than a cell which does not express said sub-lethal level of said antisense nucleic acid.

123. The method of Claim 121, further comprising:

(d) expressing a sub-lethal level of a second antisense nucleic acid complementary to a second proliferation-required nucleic acid in a second cell, wherein said second proliferation-required nucleic acid is in a different biological pathway than said proliferation-required nucleic acid in step (a); and

(e) determining whether said second cell does not have a substantially greater sensitivity to said test compound than a cell which does not express said sub-lethal level of said second antisense nucleic acid, wherein said test compound is specific for the biological pathway against which the antisense nucleic acid of step (a) acts if said second cell does not have substantially greater sensitivity to said test compound.

124. A purified or isolated nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93.

125. A compound which interacts with a gene or gene product whose activity or expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93 to inhibit proliferation.

126. A compound which interacts with a polypeptide whose expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOs.: 1-93 to inhibit proliferation.

127. A method for manufacturing an antibiotic comprising the steps of:

screening one or more candidate compounds to identify a compound that reduces the activity or level of a gene product required for proliferation, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93; and

manufacturing the compound so identified.

128. The method of Claim 127, wherein said screening step comprises performing any one of the methods of Claims 28, 38, 45, 96, 99 and 110.

129. A method for inhibiting proliferation of a microorganism in a subject comprising administering a compound that reduces the activity or level of a gene product required for proliferation of said microorganism, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a sequence selected from the group consisting of SEQ ID NOs.: 1-93 to said subject.

130. The method of Claim 129 wherein said subject is selected from the group consisting of vertebrates, mammals, avians, and human beings.

131. The method of Claim 129, wherein said gene product comprises a polypeptide comprising a sequence selected from the group consisting of SEQ ID NOs.: 298-305, 312-315, 327-353, 357-364, 372-458, 464-466, 468 and 472-479.

U.S.S.N.: 10/521,300
Filed: 1/14/05
Case No.: 21126YP

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 July 2001 (05.07.2001)

PCT

(10) International Publication Number
WO 01/48209 A3

(51) International Patent Classification[7]: C12N 15/31, 15/11, 15/10, C07K 14/245

(21) International Application Number: PCT/US00/34419

(22) International Filing Date: 19 December 2000 (19.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/173,005 23 December 1999 (23.12.1999) US

(71) Applicant *(for all designated States except US)*: **ELITRA PHARMACEUTICALS, INC.** [US/US]; Suite A, 3510 Dunhill Street, San Diego, CA 92121 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **FORSYTH, R., Allyn** [US/US]; 1135 Beryl Street, San Diego, CA 92109 (US). **OHLSEN, Kari, L.** [US/US]; 3560 Vista De La Orilla, San Diego, CA 92117 (US). **ZYSKIND, Judith, W.** [US/US]; 8415 La Jolla Scenic Drive, La Jolla, CA 92037 (US).

(74) Agent: **ISRAELSEN, Ned, A.**; 16th Floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report: 2 May 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/48209 A3

(54) Title: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION OF E. COLI

(57) **Abstract:** The sequences of nucleic acids encoding proteins required for E. coli proliferation are disclosed. The nucleic acids can also be used to screen for homologous genes that are required for proliferation in microorganisms other than E. coli. The nucleic acids can also be used to design expression vectors and secretion vectors. The nucleic acids can be used to express proteins or portions thereof, to obtain antibodies capable of specifically binding to the expressed proteins, and to use those expressed proteins as a screen to isolate candidate molecules for rational drug discovery programs. The nucleic acids of the present invention can also be used in various assay systems to screen for antimicrobial agents.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/34419

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/31 C12N15/11 C12N15/10 C07K14/245

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

SEQUENCE SEARCH, EPO-Internal, WPI Data, PAJ, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EM_PRO 'Online! EMBL; 29 January 1997 (1997-01-29) BLATTNER ET AL.: "Escherichia coli K12 MG1655 section 337 of 400 of the complete genome" retrieved from EBI, accession no. ECAE447 Database accession no. AE000447 XP002181127	5,10,13, 14, 18-21, 124
Y	the whole document	1-131
X	-& DATABASE SWALL 'Online! 1 May 1992 (1992-05-01) BURLAND ET AL.: "60 kDa inner-membrane protein" retrieved from EBI, accession no. 60IM_ECOLI Database accession no. P25714 XP002181128 the whole document	5,10,13, 14, 18-21, 124
X	-& DATABASE EM_PRO 'Online!	5,13,14,

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 November 2001	30. 11. 01

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Ceder, O

3

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	EMBL; 29 January 1997 (1997-01-29) BLATTNER ET AL.: "Escherichia coli K12 MG1655 section 298 of 400 of the complete genome" retrieved from EBI, accession no. ECAE408 Database accession no. AE000408 XP002181129 the whole document	124
A	-& BLATTNER ET AL.: "THE COMPLETE GENOME SEQUENCE OF ESCHERICHIA COLI K-12" SCIENCE, vol. 277, 5 September 1997 (1997-09-05), pages 1453-1462, XP002923023 the whole document & THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, no. 46, 13 November 1998 (1998-11-13), pages 30415-30418, & GENOMICS, vol. 16, 1993, pages 551-561, XP000615272 & GENE, vol. 38, 1985, pages 85-93, XP000915093 ---	1,9
Y	WO 99 02673 A (DUGOURD DOMINIQUE ;WRIGHT JIM A (CA); YOUNG AIPING H (CA); GENESEN) 21 January 1999 (1999-01-21) page 7, line 25 -page 9, line 30; examples 2-6 ---	1-131
X	DATABASE EM_PRO 'Online! EMBL; 19 May 1993 (1993-05-19) BURLAND ET AL.: "E. coli; the region from 81.5 to 84.5 minutes" retrieved from EBI, accession no. ECUW82 Database accession no. L10328 XP002181130 the whole document -& GENOMICS, vol. 16, 1993, pages 551-561, XP000615272 ---	5,10,13, 14,124
X	DATABASE EM_PRO 'Online! embl; 7 November 1985 (1985-11-07) CERETTI ET AL.: "Escherichia coli spc ribosomal protein operon" retrieved from EBI, accession no. ECSPC Database accession no. X01563 XP002181131 the whole document & NUCLEIC ACIDS RESEARCH, vol. 11, no. 9, 1983, pages 2599-2616, ---	5,13,14, 124

-/--

3

I ational Application No
PCT/US 00/34419

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EM_PRO 'Online! EMBL; 30 December 1994 (1994-12-30) PLUNKETT, G: "Escherichia coli K-12 chromosomal region from 67.4 to 76.0 minutes" retrieved from EBI, accession no. ECUW67 Database accession no. U18997 XP002181132 the whole document ---	5,13,14, 124
A	POST L E ET AL: "NUCLEOTIDE SEQUENCE OF THE RIBOSOMAL PROTEIN GENE CLUSTER ADJACENT TO THE GENE FOR RNA POLYMERASE SUBUNIT BETA IN ESCHERICHIA COLI" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE USA,US,NEW YORK, NY, vol. 76, no. 4, 1 April 1979 (1979-04-01), pages 1697-1701, XP000574791 abstract ---	1
A	WO 98 21366 A (Q B I ENTERPRISES LTD ;EINAT PAZ (IL); VASQUEZ IASLOP NORA CECILIA) 22 May 1998 (1998-05-22) page 8, line 9 - line 13 page 21, line 30 -page 27, line 35 ---	1
A	VAN HEESWIJK W C ET AL: "The genes of the glutamine synthetase adenylylation cascade are not regulated by nitrogen in Escherichia coli" MOLECULAR MICROBIOLOGY,BLACKWELL SCIENTIFIC, OXFORD,GB, vol. 9, no. 3, 1993, pages 443-457, XP000926027 ISSN: 0950-382X ---	
A	LEE N G ET AL: "Molecular cloning and characterization of the nontypable Haemophilus influenzae-2019 rfaE gene required for lipopolysaccharide biosynthesis" INFECTION AND IMMUNITY,AMERICAN SOCIETY FOR MICROBIOLOGY. WASHINGTON,US, vol. 63, no. 3, 1995, pages 818-824, XP000953326 ISSN: 0019-9567 ---	
A	AUSTIN A E ET AL: "Genetic analysis of lipopolysaccharide core biosynthesis by Escherichia coli k12 insertion mutagenesis of the RFA locus" JOURNAL OF BACTERIOLOGY,WASHINGTON, DC,US, vol. 172, no. 9, September 1990 (1990-09), pages 5312-5325, XP000926028 ISSN: 0021-9193 --- -/--	

3

International Application No
PCT/US 00/34419

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
T	VALVANO M A ET AL: "The rfaE gene from Escherichia coli encodes a bifunctional protein involved in biosynthesis of the lipopolysaccharide core precursor ADP-L-glycero-D-manno-heptose" JOURNAL OF BACTERIOLOGY,WASHINGTON, DC,US, vol. 182, no. 2, January 2000 (2000-01), pages 488-497, XP000926030 ISSN: 0021-9193 ———	

3

..ternational application No.
PCT/US 00/34419

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: -
because they relate to subject matter not required to be searched by this Authority, namely:

see FURTHER INFORMATION sheet PCT/ISA/210

2. [X] Claims Nos.: -
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

As a result of the prior review under R. 40.2(e) PCT, part of the additional fees are to be refunded.

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [X] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

1-131 (Seq. Id. Nos. 1, 60, 220 and 413)

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[X] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

As far as claims 129-131 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.1

Rule 39.1(iv) PCT - Method for treatment of the human or animal body by therapy

---

Continuation of Box I.2

Present claims 36, 44, 57, 95, 98, 109, and 117 relate to a compound defined by reference to a desirable characteristic or property, namely being identifiable by using the method of claims 28, 38, 45, 84, 96, 99, and 110, respectively. Present claims 125 and 126 relate to a compound defined by reference to a desirable characteristic or property, namely interacting with a gene or gene product or a polypeptide whose activity or expression is inhibited by an antisense nucleic acid comprising one of SEQ ID NOS 1-127.
The claims cover all compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to the sequences of claims 1, 9 and 19 .

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-8, 12-14, 45-124, 129-131 all partially

   Invention 1:

   A purified or isolated nucleic acid sequence, consisting of Seq Id No 1, a vector comprising said sequence, a host cell containing said vector, and their uses.

2. Claims: 1-8, 12-14, 45-124, 129-131 all partially

   Inventions 2 to 93:

   Idem as invention 1, but for Seq Id Nos 2-93, respectively.

3. Claims: 9-11, 15-44, 125-128 all partially

   Invention 94:

   A purified or isolated nucleic acid sequence consisting of Seq Id No 106, a vector comprising said sequence, a host cell containing said vector, a polypeptide encoded by said nucleic acid sequence and having Seq Id No 299, and an antibody binding said polypeptide, and their uses.

4. Claims: 9-11, 15-44, 125-128 all partially

   Inventions 95 to 231:

   Idem as invention 94, but for nucleic acid Seq Id Nos 107-112, 119-122, 134-160, 164-171, 179-265, 271-273, 275, 279-286 and corresponding polypeptide Seq Id Nos 300-305, 312-315, 327-333, 357-364, 372-458, 464-466, 468, 472-479, respectively.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 00/34419

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9902673 A	21-01-1999	AU 8327798 A	08-02-1999
		WO 9902673 A2	21-01-1999
		EP 1025219 A2	09-08-2000
WO 9821366 A	22-05-1998	AU 5442198 A	03-06-1998
		EP 0960212 A1	01-12-1999
		JP 2001504694 T	10-04-2001
		WO 9821366 A1	22-05-1998
		US 6057111 A	02-05-2000

THIS PAGE BLANK (USPTO)